The Impact of Pension Reforms on Income Inequality, Savings, and Health

Inaugural-Dissertation

zur Erlangung des akademischen Grades

eines Doktors der Wirtschaftswissenschaft

des Fachbereichs Wirtschaftswissenschaft

der Freien Universität Berlin

vorgelegt von

Stefan Etgeton, M.Sc.

aus Mettingen

Berlin, 2019

Gedruckt mit Genehmigung des Fachbereichs Wirtschaftswissenschaft
der Freien Universität Berlin

Dekan:
Prof. Dr. Dr. Andreas Löffler

Erstgutachter:
Prof. Dr. Peter Haan

Zweitgutachter:
Prof. Dr. André Decoster, Katholieke Universiteit Leuven

Tag der Disputation: 15.02.2019

Erklärung über Zusammenarbeit mit Koautoren und Vorveröffentlichungen

Kapitel 2 wurde veröffentlicht als:

Stefan Etgeton (2018). „The effect of pension reforms on old-age income inequality“. In: *Labour Economics* 53, S. 146 –161

Kapitel 3 ist bisher unveröffentlicht.

Kapitel 4 ist in Zusammenarbeit mit Anna Hammerschmid entstanden (Eigenanteil: 50%) und bisher unveröffentlicht.

Acknowledgments

First of all, I would like to sincerely and warmly thank my supervisor and boss Peter Haan. Had he not always motivated me, had he not inspired me so much, well, maybe this dissertation would have never been finished. I would also like to thank my second supervisor André Decoster, who enabled me to spent an insightful research stay in the strangest country.

I would also like to thank the FNA for supporting parts of my doctorate with a scholarship. Further, I am indebted to staff of the research data center of the Federal statistical office and statistical offices of the Länder, the research data center of the IAB, and the research data center of the DRV.

Further, DIW Berlin was an ideal place to work productively on a thesis, surrounded by good people. I am happy to have Anna Hammerschmid as a co-author. Johannes Geyer shared a lot of knowledge about social policy with me. Thore Schlaak, Sascha Drahs and Clara Welteke have made even the most strenuous days bearable. The tireless efforts of the cafeteria staff must also not go unmentioned.

Moreover, I thank my parents Georg and Gertrud for their support. I thank my brother for being my brother. I thank my sister for always being very interested in the progress of my studies. Finally, I must thank Victoria.

Table of Contents

List of Figures

List of Tables

1 Introduction

Over the last 160 years, life expectancy in the developed world increased in an extraordinarily linear manner by around 40 years (Oeppen and Vaupel 2002). Since the 1970s, fertility in most of the European countries is at levels considerably below replacement (Frejka and Sobotka 2008). The resulting demographic change is robust to any reasonable level of migration (Coleman 2008). Further, a permanent decrease of marginal tax rates of top incomes has been observed in the 1980s (Piketty and Saez 2013). In consequence, the affordability of public pay-as-you-go pension schemes is being called into question in many OECD countries.

In order to ensure financial sustainability, most countries have initiated reform processes that reduce the generosity of public pension schemes. Common policy measures include the abolition of early retirement options, a increase in the normal retirement age, and benefit reductions. The intended effects of these reforms include a prolonged work life, longer contribution periods and shorter periods in retirement. However, individual coping strategies are not limited to changes in employment. In fact, theory and the political debate highlight that private savings can alleviate the burden of pension reforms by spreading costs over the life course. Furthermore, not only the implications of pension reforms for savings and employment are of public interest. Potential impacts of reduced pension generosity on the individuals' health and wellbeing must as well be understood. Therefore, this dissertation analyzes the impact of pension reforms on employment, retirement age, income, private savings and health.

However, not all people cope equally well with pension reforms, or at least do not react in the same way to reforms. Against this background, I pay special attention to effect heterogeneities and distributional effects.

In Chapter 2, I analyze the effects of an increase of the normal retirement age (NRA) on employment outcomes. A special focus lies on the consequences for old-age income inequality. An increase of the NRA implies a financial incentive to prolong the working life and postpone retirement entry.

However, financial incentives are ineffective if jobs cannot be retained because of labor market constraints.

I estimate a structural model of labor supply decisions of elderly male workers in Germany. Subsequently, I simulate the effects of a shift of the NRA from 65 to 67 in different scenarios to draw conclusions about the evolution of employment outcomes and distributional effects. The model is identified by exogenous variation from a previous pension reform. In contrast to other studies, special attention is given to the role of involuntary job separations. Identifying the interplay of the NRA and labor market constraints as a driver of old-age income inequality is one central contribution of this study.

Labor market frictions are relevant in the retirement process. Household survey data shows that involuntary job losses are the cause for a significant share of the overall number of job exits in Germany. Every year about 6% of employed men aged 60+ involuntarily lose their jobs. In particular, low educated, sick and poor individuals, as well as those with low seniority, previous unemployment experience or foreign citizenship face a high risk of involuntarily losing their job. The chances of adequate re-employment after a job loss after turning 60 are virtually nonexistent in Germany.

Yet, Germany is not an isolated case here. Male unemployment rates of more than 7% and inactivity rates of more than 40% in, for example, Portugal, Finland, France, Spain and Greece (ages 60–64 in 2017; Eurostat 2018) are indicative that labor demand for elderly Europeans is generally limited.

I perform an *ex ante* evaluation of the shift of the NRA to age 67 that is being phased-in in Germany from 2012 to 2031 using a discrete choice model. Structural discrete choice models are particularly common in the analysis of retirement timing (see, for example Rust and Phelan 1997; Gustman and Steinmeier 2015). The model features an exogenous and individual risk of involuntary job loss that varies along socio-demographic characteristics.

The model is estimated using high quality administrative data from the German employment agency and the pension fund. Therefore, I can accurately compute accrued pension rights based on full working biographies and precise earnings information. Furthermore, the timing of employment exits and retirement benefit claims can be distinguished clearly.

From the estimation it shows that labor market frictions matter. In the main specification, my simulation suggests that the average retirement age increases by 0.6 years in response to the reform. Pension benefits decline by 2.0%. The reform has heterogeneous effects. Less educated and poor individuals are generally those who are most endangered by involuntary job loss and therefore have less possibilities to adjust their retirement timing. The frictions-caused inability to adjust retirement behavior results in increasing pension inequality. I estimate alternative scenarios varying the context of the reform: Decreasing the risk of involuntary job loss is most effective in curbing inequality. I highlight how an involuntary and unexpected job loss leads to a drop in consumption during retirement. Uniformly improving health has the opposite effect, inequality is increased.

In the structural model of Chapter 2, savings are modeled in a simplistic manner. Nevertheless, the different scenarios reveal an interesting pattern: depending on the corresponding employment effect, the effect of a pension reform on savings seems ambiguous. With respect to the first-time establishment of a pension scheme, Feldstein (1974) stresses the theoretical importance of employment effects for the overall effect on savings. In other empirical studies, decreasing pension generosity is shown to have a positive effect on private savings using benefit decreasing pension reforms (see, for example, Attanasio and Brugiavini 2003; Lachowska and Myck 2018) for identification. However, there are no empirical evaluations employing isolated changes in eligibility ages to assess the effect of pension generosity on savings. Yet, shifts of elgibility ages have employment effects as shown in Chapter 2 and studies of, for example, Mastrobuoni (2009) and Staubli and Zweimüller (2013).

Therefore, chapter 3 deals in detail with the effects of an increase of eligibility ages on savings. I make a theoretical and an empirical contribution. First, I present theoretical evidence that an increase of the early retirement age (ERA) has an ambiguous effect on private savings. I contribute by formally showing that the sign of the effect on savings rates depends crucially on the corresponding employment effect of a reform of the ERA. Further, I contribute to the literature by empirically estimating the causal effect of an increase of the ERA on private savings. Methodologically, I rely on a regression discontinuity design (RDD). Identification stems from a policy induced discontinuous jump of the ERA along neighboring female birth cohorts in Germany. In contrast to the previous literature on the effect of decreasing pension generosity on savings rates, my estimated effects are non-positive.

If employment exit age is unchanged, an increase of the ERA constitutes a loss in pension wealth. *Prima facie*, a loss of pension wealth leads to an increase in savings. However, a loss in pension wealth also increases the relative price of leisure leading to a delayed exit from employment. This delay, in return, increases pension contributions, reduces time in retirement, and again reduces the need for additional savings. Formalizing these two countervailing mechanisms and highlighting the ambiguity of the overall effect of the ERA on savings is the first contribution of Chapter 3.

The empirical part of this chapter relates to the literature on the effect of pension generosity on savings. Existing studies present evidence that benefit cuts lead to increases in private savings rates (see, for example, Attanasio and Brugiavini 2003; Lachowska and Myck 2018; Lindeboom and Montizaan 2018). In a more general setting, Chetty et al. (2014) present evidence of a low savings elasticity.

I use fine-grained household savings data to estimate the effect of an increase of the ERA on private savings rates. Identification stems from an exogenous policy variation. In 1997, an isolated reform of the ERA of German women was passed into law, taking effect in 1999. The reform increased the ERA of women discontinuously from 60 to 63. The reform, however, only affected cohorts born in 1952 or later. For women who otherwise would have retired at age 60, the reform resulted in a reduction of pension wealth between 5% and 7%. My estimation sample includes households of women born before and after the reform threshold January 1^{st}, 1952. Women are aged 45-59, that is, not yet eligible for retirement. Along treatment status, they differ in their anticipated age at employment exit and anticipated retirement age. Estimating the effect of the ERA on savings rates is a contribution to the literature.

My results show non-positive effects of an increase of the ERA on monthly private savings rates. The estimated effect of the increase of the ERA on savings rates is -1.1 percentage points. The effect is significant at the 95% confidence level and should be interpreted as an intention-to-treat-effect (ITT) because only 60% of women are affected by the reform. In a subgroup analysis, point estimates of highly educated women are more substantial (-1.5 percentage points) but no longer significant. The effect estimate in the group of low wealth households is significantly different from zero and slightly more substantial (-1.5 percentage points) than the full sample estimate.

The effect sign is in line with the anticipation of prolonged employment. In point of fact, analyzing the same reform, Geyer and Welteke (2017), find substantial effects on employment at ages 60-62 and the realized retirement age. The suggestive evidence of heterogeneity along educational groups fits the pattern found in Chetty et al. (2014), who find that less educated individuals struggle to optimally adjust to changing saving incentives.

Instead of the financial and employment dimension of pension reforms, the focus in Chapter 4 is on the health impacts of retirement. Many public debates are concerned with adequate retirement ages. However, studies for various countries and using different identification strategies come to contradicting conclusions with respect to the effects of retirement on health.

Many previous studies have used discontinuities at eligibility age thresholds as source of exogenous variation in individual retirement behavior. Other studies exploit variation from pension reforms. This chapter contributes to the literature by using variation from the strong and not gradually phased-in increase of the ERA of German women that is also used for identification in Chapter 3.

Chapter 4 contains the first study on the effect of retirement on health using variation from the German increase of the ERA of women for identification. The design of the pension reform provides a robust basis for a convincing fuzzy regression discontinuity (RDD) framework. Birth cohorts 1951 and older are unaffected, whereas the ERA of birth cohorts 1952 and younger increases from 60 to 63.

A two-sample two-stage least squares (TS2SLS) regression method is used to estimate the causal effect of retirement on health in the reform induced RDD setting. The first stage effects of the reform on retirement status are estimated using a large and precise administrative data set of the pension insurance. The second stage of the RDD model, the effect of retirement on health, is estimated using a combination of two well-established and comparable survey data sets, SOEP and SHARE. As the health outcome, self-reported health is used, a common and broad subjective health measure.

Results show that the effect of retirement on self-reported health is non-detrimental. The findings further point at effect heterogeneity along the educational dimension. Low educated women seem to benefit more from retirement, compared to the average. Point estimates do not depend on the TS2SLS method. Together with other existing studies that have found effect heterogeneity across socioeconomic groups (for example, Carrino, Glaser, and Avendano 2018; Eibich 2015), results of this chapter suggest that prolonged work lives can have aggravating effects on health inequality along socioeconomic dimensions.

The dissertation as a whole helps to understand intended and unintended effects of pension reforms. I present mild employment effects of an increase of the NRA that vary along the individual risk of involuntary job loss. Implications for income inequality are highlighted. I show that the isolated increase of the ERA does not lead to increased savings rates with effect heterogeneity along levels of wealth and education. I present evidence of a non-detrimental effect of retirement on health. The effect is stronger for less educated individuals. At large, the results of this dissertation cast doubts on whether pension reforms are socially balanced.

The following three chapters comprise the three self-contained studies on the effects of pension reforms. I discuss policy implications and conclude in Chapter 5.

2 The Effect of Pension Reforms on Old-age Income Inequality

Published as:

Stefan Etgeton (2018). “The effect of pension reforms on old-age income inequality”. In: *Labour Economics* 53, pp. 146 –161

3 The Effect of Pension Reforms on Savings Behavior

3.1 Introduction

In response to demographic change, many OECD countries are reforming their systems of old-age provision. Such reforms aim at decreasing pension generosity. A reduction of generosity should have effects on employment decisions, the realized retirement age, and private savings. Many studies analyze the effects of pension reforms on employment and retirement age empirically in a rigorous manner. In comparison, the effects of pension reforms on savings is less well studied. Yet, savings are of importance for the level of working-age and old-age consumption. In theory, the effect of pension generosity on private savings is ambiguous and crucially depends on the level of corresponding employment effects. With respect to the first-time establishment of a pension scheme, Feldstein (1974) stresses the theoretical importance of employment effects for the overall effect on savings. In empirical studies, decreasing pension generosity is shown to have a positive effect on private savings using benefit decreasing pension reforms (Attanasio and Brugiavini 2003; Attanasio and Rohwedder 2003; Feng, He, and Sato 2011; Lachowska and Myck 2018; Lindeboom and Montizaan 2018) and cross-country variation (Alessie, Angelini, and van Santen 2013) for identification. There are no empirical studies shedding light on the isolated role of pension eligibility ages for private savings. This is despite the substantial impact that pension eligibility ages have on labor supply (Mastrobuoni 2009; Staubli and Zweimüller 2013; Lalive and Staubli 2015; Geyer and Welteke 2017; Seibold 2018).

Therefore, this paper makes a theoretical and an empirical contribution. First, I present theoretical evidence that an increase of the early retirement age (ERA), in fact, has an ambiguous effect on private savings. I contribute by formally showing that the sign of the effect on savings rates depends crucially on the corresponding employment effect of a reform of the ERA.

Further, I contribute to literature by empirically estimating the causal effect of an increase of the ERA on private savings. Methodologically, I rely on a regression discontinuity design (RDD). Identification stems from a policy induced discontinuous jump of the ERA along neighboring female birth cohorts in Germany. In contrast to the previous literature on the effect of pension generosity on savings rates, my estimated effects are non-positive.

In his seminal paper, Feldstein (1974) argues that the effect of a pension scheme on savings is ambiguous. His theoretical arguments concerning the first-time establishment of a pension scheme can easily be reversed to apply to a situation of decreasing pension generosity. One example of a generosity decreasing policy is the increase of the ERA. If employment exit age is unchanged, an increase of the ERA constitutes a loss in pension wealth. *Prima facie*, a loss of pension wealth leads to an increase in savings. However, a loss in pension wealth also increases the relative price of leisure leading to a delayed exit from employment. This delay, in return, increases pension contributions and again reduces the need for additional savings. Formalizing these two countervailing mechanisms and highlighting the ambiguity of the overall effect of the ERA on savings is the first contribution of the paper.

The empirical part of my study relates to the literature on the effect of pension generosity on savings. This literature can be subdivided into two strands. One strand agrees on positive semi-structural estimates of substitutability between pension wealth and private savings (Attanasio and Brugiavini 2003; Attanasio and Rohwedder 2003; Bottazzi, Jappelli, and Padula 2006; Feng, He, and Sato 2011; Lachowska and Myck 2018). Yet, previous applications of the commonly used estimation framework do not account for employment effects in a transparent and appropriate way. Nevertheless, model identification builds on exogenous benefit cuts that are shown to imply positive and relevant employment effects (for example, Brown 2013; Manoli and Weber 2016; Bönke, Kemptner, and Lüthen 2018).

Pension scheme generosity and savings can also be related by estimating reform effects on savings rates. Although the specific interpretation is not easily generalizable, these analyses do not depend on the correct depiction of employment effects. Existing studies present evidence that benefit cuts lead to increases in private savings rates (Attanasio and Brugiavini 2003; Lachowska and Myck 2018; Lindeboom and Montizaan 2018). Lindeboom and Montizaan (2018) find small reform effects on savings but simultaneously estimate a substantial extension of employment. In a more general setting, Chetty et al. (2014) present evidence of a low savings elasticity. Using administrative data and the introduction of subsidized savings accounts, they find that low educated Danes struggle to adjust to changing saving incentives.

In the empirical part of my study, I use fine-grained household savings data to estimate the effect of an increase of the ERA on private savings rates. Identification is based on exogenous policy variation. In 1997, an isolated reform of the ERA of German women was passed into law, taking effect in 1999. The reform increased the ERA of women discontinuously from 60 to 63. The reform, however, only affected cohorts born in 1952 or later. For women who otherwise would have retired at age 60, the reform resulted in a reduction of pension wealth between 5% and 7%. My estimation sample includes households of women born before and after the reform threshold January 1^{st}, 1952. Women are aged 45-59, that is, not eligible for retirement. However, they differ along their treatment status in their anticipated age at employment exit and anticipated retirement age. Estimating the effect of the ERA on savings rates is the main contribution of this paper.

My results show non-positive effects of an increase of the ERA on monthly private savings rates for every specification and every subgroup. In the baseline specification, I include all households of women close to the birth threshold of the reform. The estimated effect of the increase of the ERA on savings rates is -1.1 percentage points. The effect is significant at the 95% confidence level and should be interpreted as an intention-to-treat-effect (ITT) because only 60% of women are affected by the reform. In a subgroup analysis, point estimates of highly educated women are more substantial (-1.5 percentage points) but no longer significant. Point estimates in the group of low wealth households are significantly different from zero and slightly more substantial (-1.5 percentage points) than the full sample estimates.

The effect sign is in line with the anticipation of prolonged employment. In point of fact, analyzing the same reform, Geyer and Welteke (2017), find substantial effects on employment at ages 60-62 and the realized retirement age. My non-positive effects can possibly be reconciled with the positive effects found by Attanasio and Brugiavini (2003), Lachowska and Myck (2018), and Lindeboom and Montizaan (2018) through the differing effects on employment of respective reforms. Whether a non-positive effect on the per period savings rate also implies a non-positive effect on lifetime savings, can only be evaluated using a more complex estimation framework incorporating imputations of employment exit and retirement age. For now, it remains an open question. The suggestive evidence for heterogeneity along educational groups fits the pattern found in Chetty et al. (2014). A large estimate in the sample of low wealth households is indicative of a high marginal utility of consumption and a, relatively, modest cost of prolonged careers in this group.

A detailed derivation of the theoretical results is found in Section 3.2. Section 3.3 describes the institutional background in Germany. The RDD methodology and the identification strategy are discussed in Section 3.4. A brief data overview is given in Section 3.5. Graphical evidence and regression results are shown in Section 3.6. Section 3.7 concludes.

3.2 Theoretical Model

The shift of the ERA can be analyzed in a small theoretical framework. It shows that the overall effect of a pension reform on the savings rate is ambiguous and depends on the corresponding employment effect.

The model builds on ideas presented first by Feldstein (1974) and Feldstein (1976). However, the line of argumentation is reversed here to accommodate a *loss* in pension generosity instead of an increase. Also, I explicitly argue along the implications of a shift of the ERA, whereas the line of thought of Feldstein is more general.

In a simple life cycle framework with perfect foresight, an individual lives for three periods, dies afterwards and has no children. Individuals are assumed to be single; that is, no intra-household transfers can be made. In periods in which the individual works, she earns a wage w, makes mandatory retirement contributions t, and privately saves the amount s. Her contributions t finance later pension benefits. The individual sum of state provided pension benefits equals former contributions to the pension scheme in case of retirement after period 1 (the initial ERA). In case of retirement after period 2, the sum of benefits equals the sum of contributions times a correction factor $\gamma < 1$. The correction factor γ mimics the actuarial unfairness we see in most pension schemes that allow for early retirement.[1] Savings s made during the work-life are another resource to be consumed during non-employment. Throughout my analysis, I assume interest rate $r_s = 0$ is applicable to savings, an interest rate $r_t = 0$[2] is applicable to pension contributions, and a discount factor $\beta = 1$, for simplicity.

I start with a baseline (non-reform) case that is characterized by only one period of employment followed by retirement at the ERA and two periods of non-employment. Retirement at the ERA is assumed to be individually optimal.

[1] Benefits are often adjusted to account for the duration of benefit receipt, but this adjustment usually is not actuarially fair. Therefore, the incentives to continue employment after reaching the ERA are limited.

[2] Results presented generalize to a Pay-as-you-go (PAYG) pension scheme, if r_t is interpreted as the rate of population and productivity growth.

Consumption in period 1 is determined by wage w of which contributions and savings are subtracted:

$$c_1 = w - t - s \tag{3.1}$$

Consumption smoothing derived from a concave utility function requires an equal division of the sum of private savings and pension benefits over consumption in periods 2 and 3. For retirement at the ERA, contributions equal the sum of pension benefits. This results in the following consumption pattern for periods 2 and 3:

$$c_2 = c_3 = \frac{t + s}{2} \tag{3.2}$$

Leisure and consumption are assumed to be non-complementary. Because of the concavity of the utility function, a discount factor $\beta = 1$ and zero interest rates, the individual wants to keep consumption constant over the course of her complete life. Accordingly, she saves an optimal amount s^* such that $c_1 = c_2 = c_3$. Using the implications of consumption smoothing to solve for s and c leads to the optimal amounts s^* and c^*:

$$s^* = \frac{2w - 3t}{3} = \frac{2}{3}w - t \tag{3.3}$$

$$\Leftrightarrow c^* = \frac{w}{3} \tag{3.4}$$

Now, a reform changes the possibilities of the individual. The reform increases the ERA by one period. The new ERA restricts the access to pension benefits to period 3. The wish to smooth individual consumption is not challenged by this reform, $r_s = r_t = 0$ and $\beta = 1$ are still valid. However, the level of savings necessary to smooth consumption in case of the reform differs from the baseline case. The optimal amount of savings varies along the employment effect of the reform. We can distinguish two scenarios: Optimal savings of individuals who do work one period *longer* and those who do *not* prolong their careers, denoted by s_l^* and s_n^*, respectively. As mentioned before, state provided pension benefits are adjusted by a factor $\gamma < 1$ if pension receipt only starts after period 2; that is, the sum of pension benefits no longer equals the sum of contributions. The factor γ reflects the actuarial unfairness embedded in many pension schemes that allow for early retirement. In other words, early retirement is financially beneficial in a net present value perspective under reasonable assumptions regarding interest rates, life expectancy, and time preferences. If it was not beneficial, there would be far less incentive to restrict early retirement.

I differentiate in my analysis of the shift of the ERA between the two possible responses to the reform in terms of labor supply. In the first scenario, I look at individuals who find it optimal to *not* prolong their career. They stop working after period 1. In period 1 they still earn a wage w, save s_n, and pay contributions t from that wage – just as in the baseline case. However, pension benefits are not yet accessible in period 2 because of the shifted ERA. This lack of pension benefits is compensated for by consuming the share ϕ of private savings. Thus, the emerging gap between work and retirement is financed solely by savings. In period 3, the remaining savings $(1-\phi)s_n$ and pension benefits are consumed. See Eqs. (3.5) to (3.7) for a formal notation.

$$c_{1-n} = w - t - s_n \tag{3.5}$$

$$c_{2-n} = \phi s_n \tag{3.6}$$

$$c_{3-n} = (1-\phi)s_n + \gamma t \tag{3.7}$$

In optimum, individuals who do prolong careers consume c_n^* each period; that is, they smooth consumption. Since individuals live for 3 periods, earn a wage w only once and lose the share $(1-\gamma)$ of their contributions t due to the actuarially not fair computation of pension benefits, their per period consumption can be denoted as

$$c_n^* = \frac{w-(1-\gamma)t}{3}. \tag{3.8}$$

Equating Eq. (3.5) with Eq. (3.6), and equating Eq. (3.6) with Eq. (3.7), we can derive two separate expressions of s_n. Equaling those two equations, we can solve for ϕ, the share of savings consumed in period 2. The choice of ϕ depends on γ, t and w,

$$\phi = \frac{w-(1-\gamma)t}{2w-(2+\gamma)t}. \tag{3.9}$$

Equating Eq. (3.8) with Eq. (3.5), we can solve for the optimal savings decisions,

$$s_n^* = \frac{2w-(2+\gamma)t}{3}. \tag{3.10}$$

Because $\gamma < 1$, it holds that $s_n^* > s^*$; that is, in the absence of an employment effect, the optimal response to the reform is to increase savings – compared to the baseline case. Again, it is important to stress that $\gamma < 1$ is not introduced through the reform of the eligibility age, but only now takes effect because we implicitly assumed the individual in the baseline case to draw benefits as early as possible. In the specific case of Germany and all else equal, $\gamma < 1$ clearly holds: a shift of pension claiming by 3 years results in slightly higher per period pension benefits upon retirement but cumulates to a substantial loss in the net present value of pension wealth of 5% to 7%.[3,4]

In a second reform scenario, I assumes a positive employment effect, the consumption pattern can be denoted as

$$c_{1-l} = c_{2-l} = w - t - s_l, \tag{3.11}$$

$$c_{3-l} = \gamma 2t + 2s_l. \tag{3.12}$$

Because the ratio of periods in employment and non-employment is flipped, optimal consumption changes considerably. After reformulation using $c_1 = c_2 = c_3$, we see that, in comparison to the baseline case, consumption rises and savings per period decline, see Eqs. (3.13) and (3.14).

$$c_l^* = \frac{2w - 2(1-\gamma)t}{3} > c^* \tag{3.13}$$

$$s_l^* = \frac{w - (1+2\gamma)t}{3} < s^* \tag{3.14}$$

If individuals extend their working life in reaction to the reform by one period, the optimal per period consumption equals a third of the sum of *two* wages minus the loss due to the actuarial unfairness of pension benefit computation. The reason behind is the reversed ratio of periods in employment and non-employment. Lifetime income from wages is increased and consumption goes up while the increased sum of contributions and a shorter period of non-employment allow for reduced per period savings.

As the main result of the theoretical section, I note that the shift of the ERA can result in higher or lower savings rates.[5] The direction of the effect depends on the corresponding employment effect, that is,

$$s_l^* < s^* < s_n^*. \tag{3.15}$$

[3]Calculations are based on an individual with 30 years of employment at the average wage level. I assume a 3% internal discount rate, account for the 3.6% per year correction factor for retirement postponement, use current life tables for Germany, and slightly vary the expected future growth rate of pension benefits.

[4]Introducing borrowing constraints or concepts of uncertainty into the model leads to similar model implications as does actuarial unfairness.

[5]Taking the lifetime perspective on savings and benefit streams, the theoretical model can easily be extended to focus on the substitutability between pension wealth and overall private savings. Under stricter assumptions concerning γ, it can be shown that even the effect of pension wealth on life-time savings is ambiguous.

Therefore, the actual effect of a shift of the ERA on the savings rate is an empirical question. Whether the employment effect of a change to the eligibility age is large enough to reduce the savings rate will be tested in the empirical part of this study.

3.3 Institutional Background

The German Statutory Pension Scheme insures 85% of the working-age population against the risk of aging. Almost all persons in dependent employment are insured with the noteworthy exception of civil servants. The pension scheme is organized as a pay-as-you-go (PAYG) system with only a negligible capital buffer. Current employees finance current retirees. Contributions to the scheme are made as a fixed percentage of the wage. Pension benefits are calculated based on pension points that are collected in accordance to contributions (and, therefore, wages) throughout the working life. Earning the year's average wage for one year results in one pension point. Few exceptions aside, earning a higher wage results in a perfectly proportional higher number of pension points. In general, the scheme does not possess re-distributive measures. Accounting for differential mortality, it therefore can be viewed as regressive (Haan, Kemptner, and Lüthen 2017). In case of retirement at the normal retirement age (NRA), accumulated pension points are multiplied with the so-called pension value (in 2018: €32.03 in West Germany, €30.69 in East Germany) to calculate monthly pension benefits. A retirement before the NRA is possible, but implies early retirement deductions of 3.6% for each year of early retirement. Deductions of 3.6% are low by international standards (Queisser and Whitehouse 2006) and not actuarially fair (Börsch-Supan et al. 2004). Consequently, many individuals retire at the earliest day possible.

Starting in the 1970s, early retirement options became available for large shares of the work force and the normal retirement age of 65 soon became meaningless. In the 1980s, men and women could retire at age 60 with generous unemployment benefits allowing for an even earlier exit from employment. However, financial pressure led to the introduction of early retirement deductions in the early 1990s. In the late 1990s an increase of the ERA from 60 to 63 was passed into law. For men, the reform process gradually shifted the ERA along birth cohorts 1946 to 1948. In contrast, the ERA of women experienced a discontinuous jump for birth cohorts 1952 and younger.[6] If eligible, women born before 1952 still have an ERA of 60.

[6]Reform details can be found in the relevant law, *Rentenreformgesetz 1999*, abbreviated as RRG 1999, announced on December 16, 1997.

Younger women face an ERA of 63. Because of the discontinuous slope in the evolution of the ERA, women are at the core of this study.

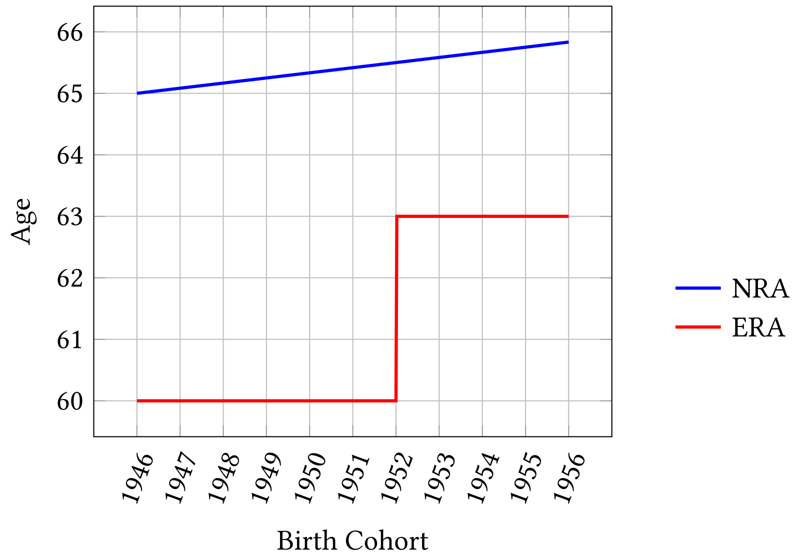

Figure 3.1: Female eligibility ages along birth cohorts

According to Geyer and Welteke (2017), before the reform, 60% of women were eligible for early retirement at age 60. Eligibility for early retirement initially required 15 years of waiting periods[7], and 10 years of obligatory contribution[8] to be acquired after age 40. The 40% ineligible for early retirement at age 60 do not experience a change of the ERA.[9] Within the group of eligible women, virtually all women face a shift of the ERA from 60 to 63.[10]

While the main policy change taking place at the birth threshold 1951/1952 concerns the ERA, the NRA is shifted at the same time by one

[7]Waiting periods are years of employment, unemployment, (up to 10) years of child rearing, and certain periods of education.

[8]Obligatory contributions are made for periods of employment, certain periods of unemployment, and (up to 3) years of child rearing.

[9]In 2018, only disabled individuals and individuals with an health-related inability to work can retire before age 63. Regulations concerning these pensions have been and still are the same for the groups to the right and to the left of the birth threshold 1951/1952.

[10]Actually, the reform changed early retirement requirements. Consequently, on average, a small subgroup of around 1% of women do not qualify for post-reform early retirement at age 63, although they would have qualified for early retirement at age 60 before the reform. For this particular subgroup, the reform even shifts the female ERA upwards from age 60 to the current NRA, that is, effectively abolishing early retirement altogether. Because of its minor empirical relevance, I ignore this group in the further course of this paper.

month for each birth cohort. As can be seen from the illustration in Figure 3.1, the shift of the NRA, however, is not discontinuous and, in comparison, of negligible size.

3.4 Methodology

The increase of the ERA only affects eligible women born after the threshold date, January 1^{st}, 1952, creating a discontinuity. In most cases, the ERA is lifted by 3 years. To the right and to the left of this birth date, the ERA is flat. Therefore, it comes as a natural choice to use an RDD to estimate the causal effect of the ERA on private monthly savings rates.

The birth date is the running variable. January 1^{st}, 1952 is the threshold date determining treatment. Because women are affected by the reform based on their birth cohort, treatment assignment can be considered exogenous. The reform, however, is only affecting women eligible for early retirement. Because not all women are eligible, the Intention-to-Treat (ITT) effect of the reform is estimated. Using linear trends, age and cohort effects are accounted for. Further, I control for other socio-demographic characteristics as education, region, homeownership and marital status. A bandwidth of 5 birth years to both sides of the birth threshold is used. I use a rectangular kernel, but results are robust to the use of a triangular kernel.

$$Y_i = \alpha + \beta X_i + \gamma D_i + \delta_1(S_i - c) + D_i\delta_2(S_i - c) + \epsilon_i \tag{3.16}$$

Equation (3.16) allows for the estimation of a causal reform effect at the reform threshold. The running variable S is defined as the birth cohort, the threshold value is set to $c = 1952$ with the treatment indicator D being defined accordingly as $D = \mathbb{1}(S \geq c)$. Further, the equation features an intercept α. Socio-demographic characteristics are denoted as X. The general cohort trend is captured by δ_1 and the diverging component of the treatment group is captured by δ_2. Outcome variable Y is defined as the savings rate.

I start by analyzing households of couples and female singles jointly. Geyer et al. (2018) show only very small spillovers of employment effects of female eligibility age on male employment outcomes. Therefore, it is reasonable to assume that estimated overall reform effects on savings are caused by the change of *female* pension wealth, employment, and earnings. In a subgroup analysis, I restrict the sample to never-married, divorced and widowed women. Zooming into this subgroup adds robustness and, furthermore, accounts for the differential importance that the reform has for couples and singles.

3.5 Data

For the empirical analysis, I employ data of the German Income and Consumption Survey (*Erwerbs- und Verbrauchsstichprobe*, EVS)[11]. The EVS is a five-yearly German household survey with the most recently available wave from 2013. It includes detailed data on consumption, income, and savings at the household level. Further, the EVS comprises socio-demographic characteristics of all household members. A fine-grained household account book is filled out by the household over three months to collect information about consumption, income, and savings. Therefore, measures of consumption, income, and savings can be considered as precise and consistent. The EVS is organized as a repeated cross section with about 60,000 households in each wave, of which 13,000 are located in East Germany. It is the only available micro data source for joint and detailed savings, wealth, and socio-demographic information in Germany. Furthermore, it is the biggest data source of its kind in Europe.

Savings rates are defined as the level of savings divided by disposable income. A sensible computation of individual-specific savings rates within the household is not possible. In general, savings cannot be assigned to a specific individual. Therefore, savings rates are computed on the household level. Most savings rates are between -0.1 and 0.3. An observational period of only 3 months is susceptible of producing extreme outliers because of durable good purchases and sales. Therefore, I trim the savings data at the 5^{th} and 96^{th} percentile.

Control variables of the later analysis include wealth, age, number of household members, as well as dummies indicating East Germans, Germans, higher education, ownership of the dwelling, widowed, divorced, and married individuals.

The estimation sample contains households of women born in years 1946 to 1956, that is the ten years surrounding the threshold date of the reform. I only use EVS waves 1998–2013 for the analysis, because earlier waves differ in terms of definitions and categorizations of savings and wealth. The reform of the ERA of women was discussed and announced in 1997. Consequentially, my sample contains no data from the pre-treatment period. The female age range for included households spans from 45 to 59 years. Thereby, households are observed before women reach the ERA — whether affected or not by the reform. Geyer et al. (2018) show that within this age group, anticipatory effects on employment are negligible. In principal, employment effects do

[11]For a short overview of the data set, see Statistische Ämter des Bundes und der Länder (2018).

Table 3.1: Summary Statistics

	Sample means	Std.dev.
Savings rate	0.111	0.158
Age (Female)	51.70	4.141
Birth cohort (Female)	1,953	2.562
Married	0.600	0.490
Widowed	0.0405	0.197
Divorced	0.210	0.407
Education (Female)	0.425	0.494
East Germany	0.256	0.436
Single	0.377	0.485
Net Income	3,655	2,202
Employed (Female)	0.677	0.468
N. of HH members	2.352	1.161
Treated	0.675	0.468
Owner of Dwelling	0.660	0.509
German	0.983	0.128
Year of survey	2,004	4.580
Observations	12,635	

Note. EVS waves 1998-2013. All. Age 45 - 59. Sample means.

not materialize before age 60. Therefore, the savings rate is not driven by diverging income and employment patterns of treatment and control group, but solely by a diverging level of savings in anticipation of a higher ERA. Furthermore, the age restriction creates, age-wise, a rather homogeneous group. Cohabiting partners of the women are restricted to be aged 40 to 60.[12] I start by analyzing couple and single households jointly; in a later step, I narrow the sample to just single women.

The final data set comprises 12,635 observations of households, among which 4,746 are female single households, that is, households of divorced, widowed, and never-married women. Summary statistics are found in Table 3.1.

3.6 Results

The results section consists of two parts. The first part graphically displays the evolution of monthly savings rates and main covariates over birth cohorts. The second part features regression results based on an RDD. Neither the graphical analysis nor the regression analysis suggest that an upward-shift of the ERA of women leads to increased savings rates.

3.6.1 Graphical Analysis

Looking at the evolution of covariates of women around the birth threshold of the ERA reform, no structural shift is detectable with the exception of age (Figure 3.2). The low frequency of the survey causes some dispersion when looking at the mean age of cohorts; see the upper left panel in Figure 3.2. It, however, is reassuring that the rate of homeownership, the share of women with higher education, the share of widowed or divorced women, the share of married women, and even female employment rates evolve rather smoothly around the threshold; see the other panels of Figure 3.2. Gray lines indicate the 95% confidence interval of the estimated linear trends. A smooth evolution of covariates around the threshold indicates comparability of individuals just born before and after the birth threshold.

The EVS is only conducted every 5 years. Therefore, cohorts differ by mean age in a systematic manner. The low survey frequency in combination with age restrictions of the sample mechanically translates into an unsteady

[12]In addition, partners (if present) are restricted to be born in years 1949 to 1956. Male cohorts 1946 to 1948 experienced a step-wise abolishment of the old age pension after unemployment or old age part-time work. As this step-wise abolishment makes them not comparable to younger cohorts of men, the exclusion is implemented. Furthermore, to deal with multiple and contradicting treatment status, I also exclude 55 homosexual couples.

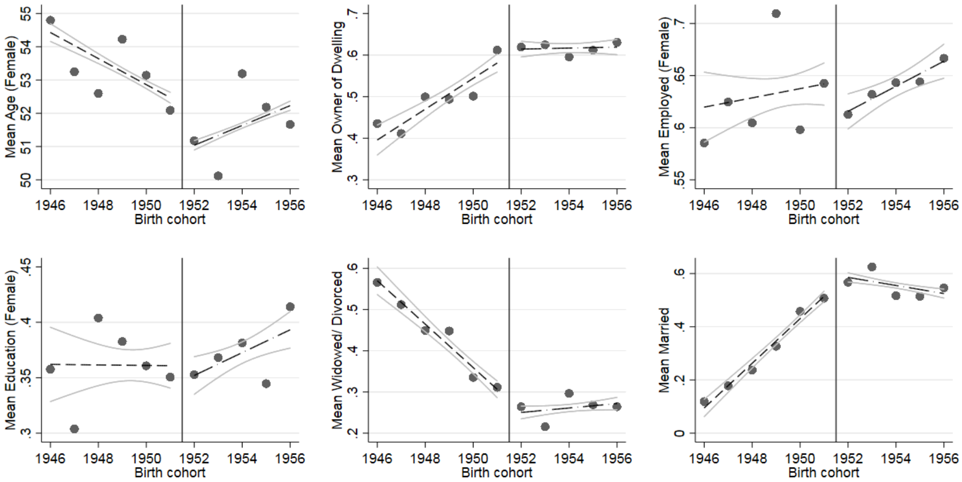

Figure 3.2: Balancing of covariates by cohort (weighted), EVS data.

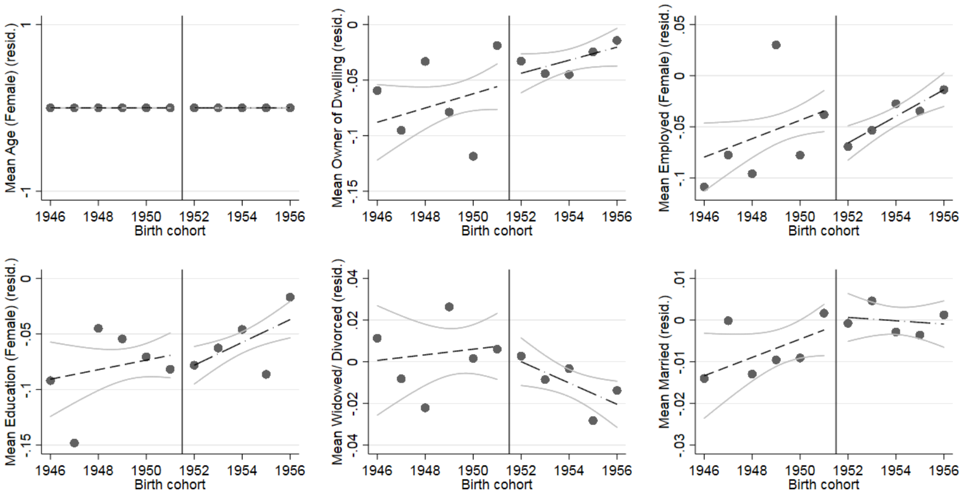

Figure 3.3: Balancing of covariates by cohort (residuals after age trend), EVS data.

and asymmetric age pattern. Furthermore, the share of couple households in the sample is higher, the older the female birth cohort. The reason is that couples are excluded from the sample if male partners are born before 1949. A restriction that leads to relatively more single households among older female birth cohorts. The reason for the restriction of male birth cohorts is the phase-out of the old-age pension for the unemployed. This type of pension was only relevant for men, and was phased-out over male birth cohorts 1946 to 1948; see Section 3.3 for details. These two technicalities in mind, I try to adjust the covariate plots. However, covariate trend patterns do not change substantially when an age trend, the impact of the household type, and a constant are subtracted; see Figure 3.3. Nevertheless, in further steps of the analysis, I account for the survey frequency based technicality of varying age patterns. The problem of over-representation of singles is explicitly tackled by repeating steps of the analysis for the sample of single households.

Moving to the comparison of savings rates, no clear differences between the affected and the unaffected groups can be detected. A graphical mean comparison of unconditional savings rates shows no positive structural break separating untreated and treated individuals; see the left panel in Figure 3.4. There is no apparent sign that individuals respond to a cut in pension generosity in the form of a shift of the ERA by increasing savings by a significant amount. In fact, the picture looks noisy.

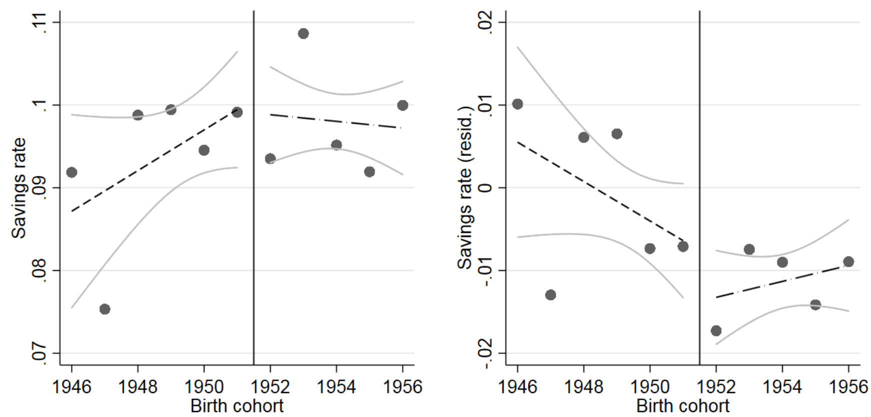

Figure 3.4: Savings rate by cohort, weighted (left) and residuals after age trend (right) — full sample, EVS data.

Again, it seems appropriate to check in how far the graphical impression is driven by varying age and household compositions. As before, I remove

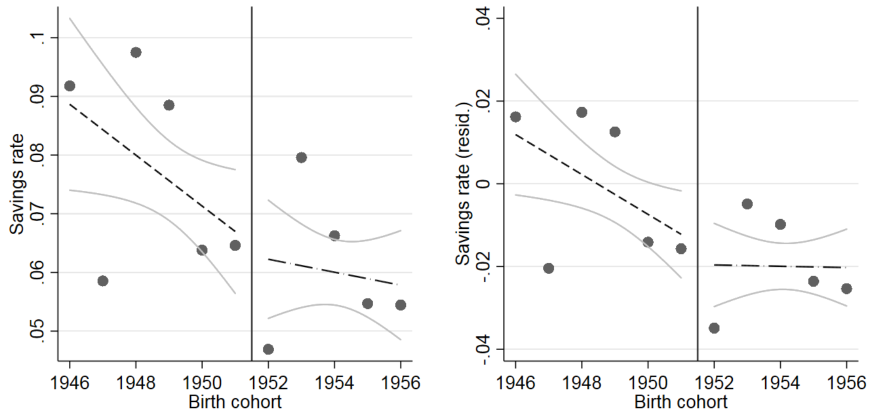

Figure 3.5: Savings rate by cohort, weighted (left) and residuals after age trend (right) — singles, EVS data.

the age trend, the impact of single households, and a constant term from savings rates. The resulting savings rate residuals are displayed in the right panel of Figure 3.4. The pattern at the threshold does not change much when compared to the unadjusted illustration. No positive effect of an increase of the ERA on the savings rate can be seen. Furthermore, a similar picture emerges when the steps are repeated while restricting the sample to single households, see the right and left panel of Figure 3.5.

3.6.2 Regression Analysis

A regression analysis allows to jointly control for a larger set of covariates, potentially reducing noise and bias further. Specifically, I control for age, cohort, homeownership, household type, marital status, region, and education. As in the graphical analysis, the birth cohort trend is allowed to differ between treated and untreated cohorts. Overall, households tend to react to an upward-shift of the ERA by decreasing monthly savings rates. The point estimate of the reform effect is -1.1 percentage points; see the first column in Table 3.2. This effect is significant at the 95% level of confidence. The confidence interval spans wide into the negative domain and the upper bound falls short of covering the zero (upper bound at -0.1 percentage points).

Running separate regressions with a stratification along female education level does result in treatment effects not significant at the 95% level. However, households of high educated women show an even larger negative point estimate of -1.5 percentage points; see Column 3 of Table 3.2. The 95%-

Table 3.2: RDD: Shift of ERA on Savings Rates, Full Sample

	(1) All	(2) Low Education	(3) High Education	(4) Low Wealth	(5) High Wealth
Treated	-0.011	-0.009	-0.015	-0.015	-0.008
	[-0.022; -0.001]	[-0.022; 0.004]	[-0.031; 0.002]	[-0.028; -0.001]	[-0.023; 0.008]
Age (Female)	-0.003	-0.004	-0.002	-0.004	-0.002
	[-0.004; -0.002]	[-0.005; -0.003]	[-0.003; -0.001]	[-0.005; -0.003]	[-0.004; -0.001]
Birth cohort (Female)	-0.001	-0.002	0.000	-0.004	0.002
	[-0.004; 0.002]	[-0.007; 0.002]	[-0.005; 0.006]	[-0.008; 0.000]	[-0.004; 0.007]
Birth cohort X Treated	0.001	0.003	-0.000	0.004	-0.002
	[-0.003; 0.005]	[-0.002; 0.008]	[-0.007; 0.006]	[-0.001; 0.009]	[-0.008; 0.004]
Other controls	✓	✓	✓	✓	✓
Observations	12,635	7,269	5,366	6,362	6,273
R^2	0.049	0.052	0.047	0.046	0.013

Note. EVS waves 1998-2013. Couples and Singles. Age 45 - 59. 95%-Confidence interval below the effects.

confidence interval ranges from negative 3.1 percentage points to positive 0.2 percentage points. For households of low educated women, the estimated effect is closer to zero and the confidence interval spans wider into the positive domain. The point estimate is -0.9 percentage points. The lower and upper bounds of the confidence interval stand at -2.2 percentage points and 0.4 percentage points, respectively; see Column 3.

The effect heterogeneity is intuitive. The employment prospects of less educated individuals are low in Germany once they reach age 55. Thus, it might be that savings behavior is heterogeneous because low educated individuals cannot as easily change their actual timing of employment exit upon facing a reform; see, for example, the findings in Chapter 2. Offering another potential explanation, Chetty et al. (2014) find that the financial illiterates have troubles benefiting from changing savings environments. In a reform setting that changes incentives to save, their savings elasticity shows to be zero.

Surprisingly, the pattern reverses when splitting the sample at the median (equivalenced) wealth. The low wealth group shows a pronounced effect of -1.5 percentage points, statistically significant at the 95% level. The high wealth group shows a non-significant and lower effect of -0.8 percentage points. It could be that, for the low wealth group, the marginal utility of consumption is high, therefore the substantive change in savings. In return, we would think, that the marginal disutility of prolonged work is rather low in the low wealth group. Consequentially, employment effects of this group should be high. Yet, I am not aware of a data set that allows for looking at the realized employment effect along wealth deciles. Nevertheless, the heterogeneity along wealth remains a bit of a puzzle, because high wealth households have, on average, higher savings rates. Thus, they actually have more leeway to downward adjust their savings rates in response to a reform. Yet, they do not.

In general, regression results are robust to the type of household the analysis is based on. Couple households should, *prima facie*, be less affected by the reform of the female ERA then should female single households. In (heterosexual) couples, one partner is unaffected by the reform. Therefore, single households are expected to react more strongly to an increase of the ERA. Yet, point estimates of single women are not distinguishable from full sample estimates and the same applies to the education subsamples; see Columns 1 to 3 in Table 3.2. In fact, single household estimates are surprisingly close the estimates derived from the full sample, but with large standard errors. The null-hypothesis that effects are identical cannot be rejected. The point estimate from the single household sample is -0.9

Table 3.3: RDD: Shift of ERA on Savings Rates, Single Households

	(1) All	(2) Low Education	(3) High Education	(4) Low Wealth	(5) High Wealth
Treated	-0.009	-0.004	-0.014	-0.005	-0.015
	[-0.026; 0.008]	[-0.027; 0.018]	[-0.041; 0.013]	[-0.026; 0.016]	[-0.043; 0.012]
Age (Female)	-0.002	-0.003	-0.001	-0.002	-0.003
	[-0.004; -0.001]	[-0.005; -0.002]	[-0.003; 0.001]	[-0.003; -0.001]	[-0.005; -0.001]
Birth cohort (Female)	-0.003	-0.005	-0.000	-0.004	-0.002
	[-0.007; 0.001]	[-0.010; 0.000]	[-0.007; 0.006]	[-0.009; 0.001]	[-0.009; 0.004]
Birth cohort X Treated	0.002	0.005	-0.001	0.003	0.001
	[-0.003; 0.008]	[-0.002; 0.013]	[-0.010; 0.008]	[-0.004; 0.011]	[-0.008; 0.011]
Other controls	✓	✓	✓	✓	✓
Observations	4,746	2,528	2,218	2,394	2,352
R^2	0.036	0.036	0.041	0.015	0.022

Note. EVS waves 1998-2013. Single women. Age 45 - 59. 95%-Confidence interval below the effects.

percentage points; see Column 1. The confidence interval spans from -2.6 to 0.8 percentage points.

Interestingly, low wealth single households react less to the reform than the high wealth group, reversing the relation we observe in the full sample. Standard errors are large and I do not want to stretch interpretation too far. However, a possible mechanism for the reversal could be the much higher labor market attachment of high wealth single women at ages 44-59 when compared to low wealth single women (+20 percentage point). The higher labor market attachment before age 60 should translate into heavier employment effects of the reform at ages 60 to 63, lowering the need for savings. In contrast, low wealth and high wealth *couples* differ in terms of female employment by only 1 percentage point.

Summing up, it should be noted that all specifications produce negative point estimates. While only two specifications yield statistically significant estimates at the 95% level, we also saw that upper bounds of the confidence intervals consistently fail to exceed levels of economic significance. Therefore, I cautiously interpret the results as evidence of non-positive effects of an increase of the ERA on savings rates. Results are indicative of more negative effects of low wealth couples and households of high educated women.

At first glance, the non-positive effects found in this study throughout various specifications are in conflict with previous studies. Attanasio and Brugiavini (2003) and Lachowska and Myck (2018) present reduced form evidence that a decrease of pension generosity substantially increases the savings rates of affected cohorts by 9 to 17 percentage points and up to 5 percentage points, respectively. Yet, the qualitative difference between the high positive estimates found by Attanasio and Brugiavini (2003) and Lachowska and Myck (2018) and non-positive results of this study could be due to the specific nature of the different pension reforms — a reduction of benefit levels vs an isolated increase of the ERA. A positive effect on savings is theoretically in line with a low employment effect, as shown in Section 3.2, and vice versa. To the best of my knowledge, no studies exist concerning the employment effects of the pension reforms analyzed by Attanasio and Brugiavini (2003) and Lachowska and Myck (2018). However, Bottazzi, Jappelli, and Padula (2006) present survey evidence that the middle-aged population analyzed by Attanasio and Brugiavini (2003) expects to *retire* on average 2.5 years later in response to the reform. This is despite the substantial benefit cuts of up to 35% implemented by the reform in question (Bottazzi, Jappelli, and Padula 2006). In comparison, the German reform of the ERA reduces the pension wealth of individuals who otherwise would retire at age 60 by 5% to 7%, all else equal. In the group of eligible women, compliers increase their retirement age by 3 years and employment is prolonged by 1.8 years[13]. Therefore, differences in the effects of Attanasio and Brugiavini (2003) and the effects estimated in this study might be rooted in varying employment effects.

In a related setting, Lindeboom and Montizaan (2018) analyze a Dutch reform that reduced pension benefits by 9%. Lindeboom and Montizaan (2018) find both positive employment effects and positive effects on the savings rate. Yet, the proportions matter. Individuals mainly repair their benefit loss by prolonging employment, on average, by 10 months. At the same time, individuals increase savings on average by an amount worth 3 months of earlier retirement, that is, increase the savings rate by 2-3 percentage points.[14] Looking at the true nature of the reform analyzed by Lindeboom and Montizaan gives hints as to why employment effects are substantial despite the relatively small cut in pension benefits. While *de jure* the reform was a reduction of pension benefits, the political debate and information letters stressed the possibility to work 13 months longer to exactly compensate for the loss in benefit levels through additional contributions and actuarial premiums. Exactly this behavioral change proved to be popular. Seen that way, the distinction between a reform of the eligibility age and a benefit cut becomes blurry, high employment effects are unsurprising, and relatively small effects on savings rates seem plausible.

3.7 Conclusion

This paper analyzes the effect of a generosity decreasing pension reform on savings rates. In particular, I evaluate how a shift of the ERA of German women affects household savings rates.

In a small theoretical model, I show that the effect of an increase of the ERA on savings rates is ambiguous if employment effects are accounted for. In the empirical part, I use German survey data to estimate reduced form effects of a reform of the ERA of women on savings rates of the household. I present evidence that an increase of the ERA of the majority of women by 3 years

[13]Own calculation based on Geyer and Welteke (2017). For the (female) population averages, these numbers have to be multiplied by the share of actual eligible women and by the share of initial compliers among the eligible women. In any case, it is obvious that the relation between loss in pension wealth and the effect on retirement age and employment exit will hardly match the Italian case as presented in Attanasio and Brugiavini (2003) and Bottazzi, Jappelli, and Padula (2006)

[14]Effect on savings rates from own calculations; based on a 70% replacement rate and a reform announcement of 6 to 7 years prior to the initial early retirement age (cf. Lindeboom and Montizaan 2018).

has non-positive effects. In the baseline specification, a 1.1 percentage point reduction of the savings rate is estimated as the ITT effect of an increase of the ERA by 3 years. The effect sign is in line with the substantial employment effects of the respective reform of the ERA (compare Geyer and Welteke 2017). The point estimate changes only slightly when the sample is restricted to households of single women.

There is indication of effect heterogeneity along levels of educational attainment and levels of wealth. Stratification along education levels leads to negative point estimates while confidence intervals partially cover small positive values. The effect of the reform seems to be pronounced if the female education level is high. Possible channels include old-age employability (see Chapter 2) and financial literacy (Chetty et al. 2014). Highly educated individuals are able to work longer in response to pension reforms. They can reduce savings more easily as working longer is a possible response to a pension reform. To the contrary, low education is associated with higher rates of involuntary job separations. This hampers attempts to prolong careers. Thus, low educated individuals might find it optimal to not alter their savings plans in response to a pension reform because planning ahead is difficult and reducing savings is risky. As an alternative explanation, Chetty et al. (2014) suggest that low education is correlated with an inability to create optimal savings plans or revise them.

Looking at the heterogeneity along levels of household wealth, we see that low wealth households show a larger negative effect on savings rates in response to the reform. A high marginal utility of consumption within this group is a potential mechanism. A high marginal utility of consumption should lead to the low levels of wealth that we observe. In response to the increase of the ERA, a high marginal utility of consumption should also lead to prolonged employment. However, the employment effects after age 60 stratified along wealth levels are not observed.

The need for private old-age provision is a prominent and recurring topic in the political debate surrounding German pension reforms of the late 1990s and the 2000s. While recommendations to boost private savings to compensate for the restrictions of statutory pension scheme generosity are numerous, poverty-vulnerable groups fail to follow recommendations. In general, the take-up of newly designed subsidized pension accounts is substantial but partakers often have middle and high incomes (Börsch-Supan et al. 2015). As shown in this study, a less generous retirement scheme might not lead to increased private savings but prolonged careers. In this light, the danger of crowding out of regular savings through subsidized savings accounts becomes apparent. German savings subsidies might be misdirected.

This is in line with Chetty et al. (2014), who show that subsidized savings accounts pose a risk of substantial crowding out.

My reduced form findings do not relate to the literature on the substitutability of pension wealth and savings directly (cf. Gale 1998; Attanasio and Brugiavini 2003; Attanasio and Rohwedder 2003; Feng, He, and Sato 2011; Lachowska and Myck 2018). While I present evidence that the effect of an increase of the ERA on savings rates is non-positive, drawing conclusions about the substitutability of pension wealth and private savings would require switching to the lifetime perspective of employment, consumption, and savings. A sound assessment of the substitutability of pension wealth and savings rates requires a semi-structural modeling framework that takes into account life-expectancy, retirement age, age of employment exit, pension entitlements, initial wealth, and earnings. Existing research does not account for employment effects when estimating the substitutability of pension wealth and savings. Yet, this and numerous other studies show that employment effects are relevant in the context of pension reforms. In this light, it could be a fruitful endeavor to check how far substitutability estimates of pension wealth and savings are sensitive to modeling decisions and assumptions. In particular, estimates might hinge on the handling of employment effects and corresponding gains in lifetime earnings and pension contributions. This task, however, is left as an avenue for future research.

4 The Effect of Retirement on Health: Evidence from a German Pension Reform

4.1 Introduction

Demographic change, in particular driven by increasing life expectancy and low fertility rates, is of growing importance in many countries. Therefore, many debates are concerned with adequate retirement ages, with numerous countries already having raised retirement entry ages in the past years (see, for example, OECD 2017) to counteract and balance the intergenerational contract of pay-as-you-go pension systems. It is, however, important to acknowledge that economic sustainability is not the only dimension that matters from a societal point of view. Potential impacts of prolonged work lives on the individuals' health and wellbeing must also be understood and taken into account when deciding on the parameters of an economically and socially suitable retirement system. However, studies for various countries and using different identification strategies come to contradicting conclusions.[1]

Our paper contributes to a better understanding of the causal effects of retirement on health. Many existing studies use discontinuities at age eligibility thresholds as source of exogenous variation in individual retirement behavior (for example, Rohwedder and Willis 2010; Eibich 2015). Other studies exploit variation from pension reforms (for example Charles 2004; De Grip, Lindeboom, and Montizaan 2011; Bloemen, Hochguertel, and Zweerink

[1]For studies that find a positive effect of retirement or a negative effect of prolonged work lives, see, for example, Charles (2004), Coe and Zamarro (2011), De Grip, Lindeboom, and Montizaan (2011), Bloemen, Hochguertel, and Zweerink (2013), Eibich (2015), Leimer (2017), Blake and Garrouste (2017), Kolodziej and Garcia-Gomez (2017), and Carrino, Glaser, and Avendano (2018). Negative effects are found by, among others, Rohwedder and Willis (2010), Kuhn et al. (2018), Godard (2016), and Fitzpatrick and Moore (2016).

2013; Blake and Garrouste 2017). Our study contributes to this literature by using variation from an unusually strong and immediate, that is, not gradually phased-in, pension reform.

To the best of our knowledge, this is the first paper on the effect of retirement on health that uses variation from the German 1999 pension reform for identification. The reform is particularly well suited as identifying variation for two reasons. First, the reform increases the ERA of women strongly, by 3 years, from 60 to 63. Second, the reform is not phased-in stepwise but takes full effect starting with birth cohort 1952. Thus, it creates a substantial discontinuity between the cohorts born in 1951 and 1952. The design of the pension reform provides a robust basis for a convincing fuzzy regression discontinuity (RDD) framework. The impact of the reform on labor market outcomes is robust and strong (Geyer and Welteke 2017; Geyer et al. 2018).

We employ a two-sample two-stage least squares (TS2SLS) regression method to estimate the causal effect of retirement on health in the reform induced RDD setting. Exploiting a cohort-based pension reform to measure effects on health outcomes requires a large sample to enable a quite narrow bandwidth around the reform cutoff for estimation of the RDD model. Therefore, we estimate our first stage effects of the reform on retirement status using a large and precise administrative data set of the pension insurance with observations of nearly 4% of the German population. Furthermore, the second stage of the RDD model, the effect of retirement on health, is estimated using a combination of two well-established survey data sets. The main survey data set is the Socio-Economic Panel Study (SOEP). To increase sample size, we complement the data set with the German sample of the Survey of Health, Ageing, and Retirement in Europe (SHARE). In order to obtain a comprehensive picture of the health effects of retirement, we analyze self-reported health, a common and broad subjective health measure that is available in both surveys. In fact, self-reported health is an oft used measure in the literature that enables comparability of our results to a broad range of other studies.

Our results suggest that the effect of retirement on self-reported health is non-detrimental. Our findings further point at effect heterogeneity along the educational dimension. Low educated women seem to benefit more from retirement, compared to the average. These results are robust to a range of sensitivity analyses. Most importantly, they are not confounded by the compulsory school reform that affected some of the cohorts in our analyses. Moreover, the conclusions do not depend on the TS2SLS method being used. Together with other existing studies that find effect heterogeneity across socioeconomic groups (for example, Carrino, Glaser, and Avendano 2018; Eibich 2015), our insights suggest that prolonged work lives can have aggravating effects on health inequality along socioeconomic dimensions.

The paper is structured as follows. Section 4.2 presents the German pension system and the 1999 pension reform, which provides the basis for our empirical approach. Section 4.3 discusses the empirical strategy and challenges for identification. Thereafter, we introduce the data sets in Section 4.4. Results are presented in Section 4.5. Section 4.6 concludes.

4.2 Institutional background

The German statutory pension scheme (Gesetzliche Rentenversicherung) is a pay-as-you-go scheme. Participation is mandatory for most workers with the notable exceptions of civil servants (who have a separate system of old-age provision) and the self-employed. Around 85% of the workforce in Germany are insured in the statutory pension scheme (Börsch-Supan and Wilke 2004). The scheme features only a few redistributive elements. It is characterized by a strong link between the individual's lifetime earnings and the benefit amount. Accounting for differential mortality, the system can be characterized as regressive (Haan, Kemptner, and Lüthen 2017).

Next to the statutory pension scheme, there are also occupational and private pension plans. These are of minor (but growing) importance – both in terms of benefit level and in the number of entitled individuals. However, less educated, less financially educated, and low income households lag behind in terms of the spread of occupational and private pension products (Börsch-Supan et al. 2015).

4.2.1 The abolishment of the old-age pension for women

The old-age pension for women (Altersrente für Frauen) granted women the possibility to retire early at the age of 60. Eligibility required a waiting time in the pension scheme of at least 15 years. All periods of employment, unemployment, and child rearing counted toward this waiting time. Furthermore, eligibility required 10 years of active contribution after the age of 40. Periods of active contribution include employment and short-term unemployment.

With the 1999 pension reform, however, the old-age pension for women was abolished for women born on or after January 1, 1952; see the relevant law, Rentenreformgesetz 1999 (1997). For these women, the earliest possible retirement age effectively increased to 63, the age at which the old-age pension for long-term insured is accessible.[2] To be eligible for a pension at the new ERA of 63 years, the affected cohorts born after January 1, 1952, must meet a different criterion. Instead of 15 years of waiting time and 10 years of active contribution, the minimum waiting time is 35 years with no further requirements. Geyer and Welteke (2017) show that these two different eligibility rules eventually turn out to cover almost identical groups. Around 60% of the women born in 1951 are eligible for the old-age pension for women with age 60. Around 59% of women born in 1952 are eligible for the pension for long-term insured individuals. Individuals who meet the new regulation for early retirement at age 63, are eligible for early retirement at age 60 under the old regulation in almost all cases — and vice versa.

As of 2014, the old-age pension for women loses relevance because the youngest eligible women, who not claimed a pension yet, turn 63, thereby reaching the ERA required for access to the pension for the long-term insured. As of 2017, the old-age pension for women is effectively abolished: the youngest potentially eligible women, those born in 1951, reach their cohort's normal retirement age (NRA), 65 years and 5 months, but access to the old-age pension for women is only possible before the NRA.

4.3 Empirical strategy

4.3.1 Challenges for identification

The raw correlation between retirement and health is potentially biased and does not necessarily reflect the causal effect of the individual's retirement status on health. There are, in particular, three potential biases being discussed in the literature on retirement and health: omitted variables bias, simultaneity, and justification bias (see, for example, Eibich 2015).

The former two biases can be tackled exploiting exogenous variation in the retirement status. Therefore, we use a fuzzy Regression Discontinuity Design (RDD) (see, for example, Imbens and Lemieux 2008; Lee and Lemieux 2010, and references therein) on the pension reform of 1999 in Germany. This reform causes a discontinuity in the retirement probability at the birth threshold January 1, 1952. In a fuzzy RDD, we use this discontinuity to instrument the individual retirement decision. Thereby, we only rely on

[2] As depicted in Table 4 in Geyer and Welteke (2017), only invalidity and disability pensions are generally available before age 63. Yet, Geyer and Welteke argue that program substitution into the invalidity pension is unlikely because of lower deductions in the invalidity scheme. Moreover, there is virtually no substitution into the disability program.

the variation in retirement that is driven by the reform induced change in retirement rules.

Justification bias exists if retired individuals report a worsening of the individual health status to justify why they are not working. Justification bias cannot be directly tackled by our empirical approach. The only way to rule out justification bias is to use objectively measured health instead of subjective assessments. The Body-mass-index (BMI) is such an objective health measure observed in our data. Yet, BMI is not expected to react in the very short run and is only observed biannually. Therefore, in our analysis of the short-run health effects of retirement we do not use BMI. For our subjective outcome variable self-reported health, justification bias would mean a downward bias of our estimates (see also, for example, Eibich 2015).

4.3.2 Two-sample two-stage least squares (TS2SLS)

We use a Regression Discontinuity framework in the context of the 1999 pension reform for identification of the causal effect of retirement on health. More specifically, our empirical strategy is a fuzzy RDD because the month of birth does not perfectly determine retirement eligibility, and, consequently, does not perfectly determine retirement status.

We employ a two-stage least squares (2SLS) approach to estimate the fuzzy RDD model. Using the two-sample version of the common two-stage least squares estimator (TS2SLS), we exploit the advantages of survey data and administrative records. Our data set for the second stage is a combination of the SOEP and SHARE survey data. In the first stage, we use administrative records from the pension fund (see Section 4.4 for details).

The TS2SLS was first proposed by Klevmarken (1982). A related two-sample IV estimator is proposed by Angrist and Krueger (1992) and Arellano and Meghir (1992). Unlike in the one-sample case, the 2SLS and the IV estimator are numerically distinct in the two-sample case. In fact, the computationally convenient TS2SLS is more asymptotically efficient than the two-sample IV estimator (Inoue and Solon 2010).

Our estimation equations closely correspond to the standard fuzzy RDD estimated by 2SLS. The retirement status is estimated using the first stage equation,

$$r_j = \gamma_0 + \gamma_1 \mathbb{1}[c_j \geq 0] + \gamma_2 c_j + \gamma_3 c_j \mathbb{1}[c_j \geq 0] + \Gamma X_j + \nu_j. \qquad (4.1)$$

Retirement is denoted as r_j. The running variable birth cohort is defined as c_j, normalized to 0 for birth month January, 1952; positive values denote post-reform and negative values denote pre-reform cohorts. Consequently,

$\mathbb{1}[c_j \geq 0]$ is the treatment indicator. X_j is a vector of covariates. The cohort trend is allowed to break at the reform threshold.

The health effects of retirement are estimated using the following second stage equation,

$$y_i = \beta_0 + \beta_1 \widehat{r_i} + \beta_2 c_i + \beta_3 c_i \mathbb{1}[c_i \geq 0] + \Theta X_i + \epsilon_i. \tag{4.2}$$

The health outcome is denoted as y_i. Retirement status is predicted as $\widehat{r_i}$. The cohort trend is allowed to break at the reform threshold. The estimated effect of retirement on health, $\widehat{\beta_1}$, is the estimate of main interest.

Importantly, using the TS2SLS estimator, the first stage equation (4.1) is estimated on a different data set than the second stage equation (4.2). Therefore, in the notation, we use subscripts j for observations from administrative data used to estimate the first stage and subscript i for the survey data used to estimate the second stage. The prediction of the retirement status, $\widehat{r_i}$, is calculated within the survey data with the coefficients estimated from administrative data on the first stage. This is possible because covariates and instruments are observed in both data sets. Both stages are estimated using OLS. We use a rectangular kernel and a bandwidth of 2 years on each side of the discontinuity.

Standard errors are clustered on the month of birth. However, standard errors cannot be computed using bootstrapping methods because special security regulations apply to the usage of the administrative data set. In particular, it is not possible to use survey and administrative data on the same computer. Therefore, we use a cluster-robust variance estimator for TS2SLS estimation that builds on the analytic expression of the TS2SLS heteroskedasticity-robust variance estimator proposed by Pacini and Windmeijer (2016).

Our cluster-robust version of the variance estimator proofs valid in a Monte Carlo simulation exercise. Using 10,000 replications and synthetic data, the standard deviation of the coefficient estimates is similar to the mean standard errors as derived from the analytical expression; see the Appendix for details.

In the specific case of Germany and in relation to the specific reform used for identification, one potential threat for identification remains to be addressed. Schooling reforms in West Germany raised compulsory schooling from 8 to 9 years. Concerning our observational window of cohorts 1950 to 1953, four large West German federal states (North Rhine-Westphalia, Hesse, Rhineland-Palatinate, Baden-Württemberg) changed compulsory schooling within cohort 1953. Bavaria only increased the compulsory school age for the 1955 cohort, so that our sample is not affected. Other states introduced reforms much earlier. In those states, our sample cohorts 1950 to 1953 are not differently affected. Reforms of the compulsory schooling age affect health outcomes positively (see, for example, Kemptner, Jürges, and Reinhold 2011). However, effects are found primarily in men. Therefore, our estimates should not be strongly influenced by these reforms. Moreover, a positive health effect of the schooling reform means that a beneficial effect of retirement is to be interpreted as a conservative estimate. To examine whether, and to what extent, our main findings are biased, we perform a robustness check including a dummy into our set of control variables that indicates whether an individual was born in 1953 and lives now in a federal state that implemented the reform for cohorts 1953+.[3]

4.4 Data

4.4.1 Data sources and sample selection

In our empirical analysis, we employ both survey and administrative data. Self-reported health is observed in the survey data. The administrative data allow for a precise first stage estimation. The main survey data source is the German Socio-Economic Panel Study (SOEP), waves 2010 to 2016. The SOEP is a representative annual German household panel survey. Each year, around 30,000 individuals from about 11,000 households are interviewed.[4] The SOEP regularly includes a self-reported health measure.

In addition to the SOEP, we use data of the Survey of Health, Ageing, and Retirement in Europe (SHARE) to obtain a larger sample for the second stage.[5] SHARE is a survey of populations aged 50 and up that is conducted in many European countries. We restrict our attention to the German sample of waves 4 (2011/12), 5 (2013), and 6 (2015). The size of the German SHARE

[3]Since we do not observe in which federal state the individuals went to school, we take the state where the individual lives today as an approximation, assuming the absence of selective migration.

[4]For more information about the SOEP, see Goebel et al. (2018). This paper uses the version 33.1 of the SOEP as described under doi:10.5684/soep.v33.1.

[5]This paper uses data from SHARE Waves 4, 5, and 6 (10.6103/SHARE.w4.600, 10.6103/SHARE.w5.600, 10.6103/SHARE.w6.600), see Börsch-Supan et al. (2013) for methodological details. The SHARE data collection has been primarily funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: N°211909, SHARE-LEAP: N°227822, SHARE M4: N°261982). Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

sample ranges between 1,500 and 6,000 observations across these three waves (Malter and Börsch-Supan 2015; Malter and Börsch-Supan 2017).

We use both SOEP and SHARE to increase statistical power in particular for heterogeneity analyses. We describe slight differences in question wording in more detail later in this Section.

To obtain precisely estimated first stage coefficients, we use the administrative VSKT data set of the German statutory pension scheme. The VSKT data is drawn as a stratified random sample from the publicly insured population. The VSKT includes 4% of the insured population.[6] Our analysis is based on the 2016 version of the data (Deutsche Rentenversicherung Bund 2016). Because the VSKT is routine data, it does not suffer from individual recall errors. Yet, the availability of socio-demographic control variables is limited. Most importantly, the household context is unobserved.

We concentrate our analysis on individuals born close to the reform implementation threshold 1951/1952. Therefore, we restrict our sample to female birth cohorts 1950 to 1953. This applies to SOEP, SHARE and VSKT alike. To identify the causal effect of retirement on health, we exploit the increase of the ERA from 60 to 63 due to the 1999 pension reform. Consequently, we only include the age range 60–62 into our analysis. Thus, we use female observations of the aforementioned cohorts from SOEP waves 2010 to 2016, and the respective observations from SHARE waves 2011/12, 2013, and 2015. In the VSKT, we use observations of the years 2010 to 2016. Accounting for missing information in both outcome and control variables, this leaves us with a maximum of 2,361 observations from SOEP and 533 observations from SHARE for an analysis of self-reported health; in total 2,894 observations. The VSKT sample is a balanced panel and contains 607,104 monthly observations of ages 60 to 62 from 16,864 unique individuals of the respective cohorts.[7] The representation of birth cohorts is roughly uniform in all three data sources; see Table 4.1. Only the uneven cohorts in the SHARE data are slightly larger. Every combination of age and birth cohort is well

[6]The 4% sample is not available as a scientific use-file and must be accessed on-site. Around 85% of the German population is insured by the pension scheme. Relevant exceptions include self-employed and civil servants. The 85% insured in the public scheme are still representative with respect to the income distribution of the German population (Bönke, Corneo, and Lüthen 2015).

[7]In general, the VSKT data set we use is comparable to the data of Geyer and Welteke (2017), who estimate the effects of the ERA reform on labor force participation and retirement status. Unlike them, we do not exclude severely disabled women, women ineligible for early retirement, and women insured by the *Knappschaft*, a subscheme of the pension fund for workers in the mining industry. Not excluding these groups makes the underlying population of the administrative data and survey data comparable. Using a newer wave of the data, we can extend our analysis by one year and also include women aged 62.

represented in the survey data (more than 100 observations per combination; not shown).

4.4.2 Main variables

Retirement is the main explanatory variable of our analysis. Retirement is defined as an indicator taking the value one if an individual receives retirement benefits as measured in the VSKT. The labor force status is also observed in SOEP and SHARE. However, this information is self-reported and, therefore, might suffer from justification bias or recall error. Yet, in a robustness check, we use only survey information and define retirement as an indicator taking the value one if the self-reported labor force status is retirement, otherwise zero.

Table 4.1: Frequencies by year of birth and data source

	Administrative Data	Survey Data	
	VSKT	SHARE	SOEP
1950	145,008	119	568
1951	151,812	153	625
1952	152,856	116	577
1953	157,428	145	591
Subtotal	607,104	533	2,361
Total	607,104	2,894	
Uniq. Individ.	16,864	1,326	

Note. Number of available observations by year and data source conditional on all control variables being non-missing. Women Age 60 to 62 of cohorts 1950 to 1953. Monthly administrative observations based on VSKT2016 (FDZ-RV). Yearly survey observations from SOEP 2010 to 2016 and SHARE waves 4 to 6.

In the VSKT data, we observe whether individuals have a record of pension entitlements from attendance of an institution of tertiary education (usually a university or a university of applied science). If entitlements from tertiary education exist, we define the binary variable for higher education to be one, otherwise zero. In the survey data, we define higher education as having obtained degrees of category 5 or higher on the ISCED 1997 scale, which includes Bachelor's or higher degrees from a university or a university of applied science.

Health can be measured along multiple dimensions. We use the self-reported general health status, which is measured in both the SOEP and the SHARE on a 5-point scale with lower values indicating better health. Self-reported general health measures are good predictors of using physician services (see, for example, Miilunpalo et al. 1997) and mortality (see, for example, Idler and Benyamini 1997). Unfortunately, the SHARE captures self-reported health on the US scale of this widely used item whereas the European scale is used in the SOEP. Therefore, the wording of the scale is slightly different in the two surveys.[8] We address this issue in a robustness check.

Table 4.2: Summary statistics, SOEP and SHARE

	SOEP			SHARE		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Retirement	2361	0.25	0.43	538	0.22	0.41
Born 1952 or later	2361	0.49	0.50	538	0.49	0.50
Full yrs of age	2361	60.98	0.82	538	61.14	0.77
Year of Birth + Mon./12	2361	1951.95	1.15	538	1951.99	1.09
East Germany	2361	0.25	0.43	533	0.22	0.41
High Education	2334	0.30	0.46	533	0.25	0.43
Married	2361	0.73	0.44	534	0.80	0.40
School Reform	2361	0.53	0.50	0	—	—
Self-reported Health	2361	2.87	0.93	538	3.10	1.02

Note. Socio-demographics of women of cohorts 1950 to 1953, ages 60 to 62. SOEP 2010 to 2016 and SHARE waves 4 to 6. Multiple observations per individual.

4.4.3 Descriptive statistics

Table 4.2 shows descriptive statistics separately for observations from SOEP and SHARE. In the SOEP subsample, 25% of observations are retired. At 22%, this number is slightly lower in the SHARE subsample. The average age amounts to around 61 years of age in both surveys. Women in the SHARE subsample are, on average, less educated and more often married than in the SOEP. SHARE respondents report a slightly worse health condition. The fraction of observations affected by the reform of the ERA, that is, observations of women born 1952 or later, is identical in the SHARE and SOEP samples (49%). A school reform dummy shows whether an observation was

[8]In English, the US scale has the categories "excellent", "very good", "good", "fair" and "poor". The European scale ranges from "very good", "good", "fair", and "bad" to "very bad".

presumably affected by the extension of the years of compulsory schooling in the 1950s. The reform was introduced at different points in time in different states of West Germany. Because states are unobserved in SHARE, this variable is only computed for SOEP observations.

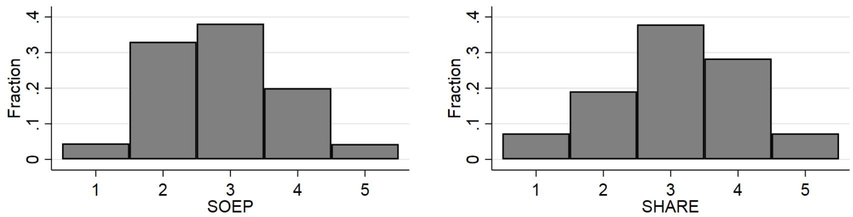

Figure 4.1: Distribution of self-reported health, SOEP 2010 to 2016 and SHARE waves 4 to 6.

Table 4.3: Summary statistics, VSKT

	N	Mean	Std. Dev.
Retirement	607,104	0.20	0.40
Born 1952 or later	607,104	0.51	0.50
Year of Birth + Month/12	607,104	1951.96	1.15
Full Years of Age	607,104	61.00	1.00
East Germany	607,104	0.23	0.42
High Education	607,104	0.23	0.42
Number of Children	607,104	1.63	1.24

Note. Socio-demographics of women of cohorts 1950 to 1953, ages 60 to 62. VSKT2016 (FDZ-RV). Unweighted, 36 observations per individual.

To obtain a more detailed picture of the comparability of self-reported health between the two surveys, we present its distributions in Figure 4.1. The distributions only look slightly different. The SOEP respondents report category 2 more often and category 4 less often, compared to the SHARE sample. This difference is probably due to differences in the wording of the categories.

Looking at the summary statistics of the administrative VSKT data (Table 4.3), 20% of the observations are retired and 51% of individuals are born after 1951. Because administrative data does not suffer from panel attrition, all individuals are observed at all ages. Therefore, the mean of completed years of age is exactly 61. Shares of highly educated individuals from administrative and survey data can be reconciled by applying sampling weights (not shown).

4.5 Results

In this section, we present our estimation results of the causal effect of retirement on self-reported health. We discuss the first stage effects in the first subsection. Thereafter, we describe the causal effects of retirement on health obtained from the second stage of our TS2SLS estimation. We provide estimated average effects as well as specific effects for women without university education to account for potential effect heterogeneity. In the third subsection, we show the robustness of our results in a range of sensitivity checks.

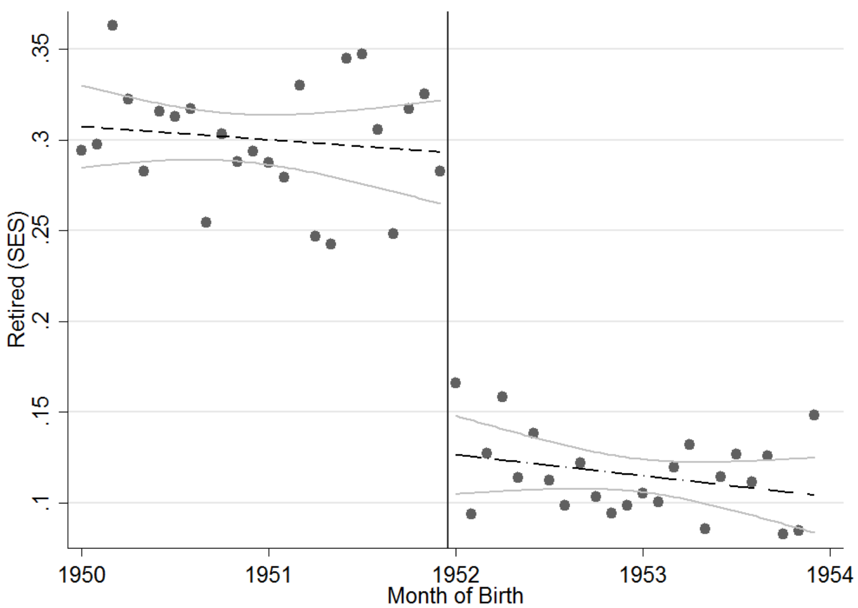

Figure 4.2: Share of retired along month of birth. Women of cohorts 1950 to 1953, ages 60 to 62. VSKT2016 (FDZ-RV).

4.5.1 First stage

The first stage – the effect of the pension reform on the retirement status – boils down to a standard fuzzy RDD problem. Detailed first stage reform analyses include Geyer and Welteke (2017), who use similar administrative data[9], and Geyer et al. (2018), who focus on the household context using the Microcensus.[10] Figure 4.2 shows the bivariate first stage relationship graphically. At the cutoff of the reform, January 1952, we observe a considerable drop in the retirement probability of more than 15 percentage points. This result is very similar to the graphical result in Geyer et al. (2018). Slight deviations in levels before and after the cutoff are most likely due to differences in the definition of retirement because of the different data sources used.

Table 4.4: First stage results

	Full Sample		Low Educ.	
	Coeff.	p-Value	Coeff.	p-Value
Treated	-0.153	0.000	-0.160	0.000
Birth Cohort	-0.009	0.318	-0.004	0.664
Birth Cohort X Treated	-0.005	0.677	-0.008	0.546
East Germany	0.165	0.000	0.190	0.000
Age in Months	0.003	0.000	0.003	0.000
Constant	-1.795	0.000	-1.730	0.000
N	607,104		487,404	

Note. Women Age 60 to 62 of cohorts 1950 to 1953 from administrative data, VSKT2016 (FDZ-RV). Standard errors clustered on month of birth. Low education category includes all individuals without any university attendance. 36 observations per unique individual.

Table 4.4 shows the multivariate estimation results of the first stage, estimated on administrative data of the public pension insurance. In the full sample, the reform causes a drop in the retirement rate of about 15 percentage points for 60 to 62 year old women, which is in line with the graphical evidence. Restricting the sample to women without higher education, we find a slightly larger drop in the retirement rate of 16 percentage points. The reform effects are estimated with great precision, which confirms the relevance of the 1999 pension reform as an instrumental variable for individual retirement. Our estimates are very close to the results of Geyer et al. (2018) who find, on average, a 16 percentage point drop in the retirement rate

[9]We are very grateful to Johannes Geyer and Clara Welteke for sharing their code with us, which we used in adapted form for the data preparation.

[10]The Microcensus is an annual obligatory survey of a random sample of 1% of the population.

and a slightly larger effect for women with lower educational background. Compared to Geyer and Welteke (2017), the ITT effect of the reform on retirement is slightly larger in our case. As mentioned in Section 4.4, Geyer and Welteke apply a stricter sample selection and focus on the age group 60 to 61 instead of 60 to 62.

4.5.2 Second stage

Table 4.5 shows the second stage results for the full sample; that is, the estimated causal effect of retirement on self-reported health. At first, the point estimate of -0.816 on a 5-point scale suggests a quite large beneficial average effect of retirement on health.[11] However, the effect is statistically insignificant. The imprecision of the estimation leads us to a cautious interpretation of the point estimate. Although the first stage of our two-stage estimation procedure is extremely precise, the second stage survey sample is probably too small to precisely pin down the causal effect of retirement on health. Having this caveat in mind, we can nevertheless rule out economically significant detrimental effects of retirement on self-reported health. The 90% confidence interval around the point estimate covers only small detrimental effects (up to values of 0.021) of retirement on self-reported health.

Table 4.5: Second stage results, full sample

	Self-rep. Health	p-Value	[90% CI]	
Retirement	-0.816	0.109	-1.652	0.021
Birth Cohort	-0.062	0.152	-0.133	0.009
Birth Cohort X Treated	0.026	0.700	-0.085	0.137
East Germany	0.095	0.325	-0.064	0.255
Age in Months	0.004	0.029	0.001	0.008
Constant	-0.271	0.851	-2.637	2.096
N	2894			

Note. Women Age 60-62 of cohorts 1950-1953 from SOEP 2010 to 2016 and SHARE waves 4 to 6 data. Standard errors clustered on month of birth. First stage estimates from VSKT2016 (FDZ-RV), years 2010 to 2016.

Turning to the subsample of women without university education, the beneficial effect of retirement on health is even larger; see Table 4.6. In comparison to the full sample, women without higher education tend to benefit more from retirement. In this subgroup, the estimated effect increases

[11] Self-reported health is coded in a way that higher values correspond to worse health.

Table 4.6: Second stage results, low education

	Self-rep. Health	p-Value	[90% CI]	
Retirement	-1.252	0.055	-2.326	-0.178
Birth Cohort	-0.050	0.307	-0.131	0.031
Birth Cohort X Treated	-0.001	0.987	-0.138	0.135
East Germany	0.196	0.131	-0.018	0.410
Age in Months	0.006	0.025	0.002	0.010
Constant	-1.132	0.532	-4.113	1.849
N	2047			

Note. Women Age 60-62 of cohorts 1950-1953 from SOEP 2010 to 2016 and SHARE waves 4 to 6 data. Standard errors clustered on month of birth. First stage estimates from VSKT2016 (FDZ-RV), years 2010 to 2016.

by about half and turns significant with a p-value of 0.055. For this subgroup of women, detrimental effects of retirement on self-reported health are unlikely. Even the upper bound of the 90% confidence interval suggests a (minimum beneficial) effect of -0.178, which roughly equals one fifth of a standard deviation of the outcome variable.[12]

Overall, the estimated effects of covariates are in line with expectations. The cohort trend is insignificant, which is in line with a quite narrow bandwidth of only two birth cohorts to each side of the reform threshold. Moreover, the trend does not break at the reform threshold. East Germans show lower self-reported health than their West German counterparts in our sample, but the coefficients are not significant. Age is associated with a detrimental effect on health.

Altogether, women with lower education seem to benefit more from retirement in terms of self-reported health compared to their higher educated

[12]To put our point estimates into context, estimates can be expressed as standard deviations of the outcome variable. In our survey sample, the standard deviation of self-reported health is 0.96. Therefore, the effects of our main specification on the sample of low educated women and the full sample have a size of slightly above and slightly below one standard deviation, respectively. Compared to the general health economics literature, these are large effects. Using lottery prizes for identification, Lindahl (2005) estimates an increase of average income of 200-500% to be necessary to cause his index of general health to improve by one standard deviation. Using SOEP survey data and identification from plant closures, Schmitz (2011) finds small and insignificant effects of involuntary unemployment on a general measure of health. His confidence intervals rule out effects of the size of our estimated point estimates. In a simple mean comparison using a general self-reported health indicator with four categories, the unconditional difference between individuals with symptoms of depression and individuals without symptoms is only 0.35 standard deviations (Mulsant, Ganguli, and Seaberg 2015).

counterparts. Possibly, jobs of lower educated women are more harmful to health. In return, retirement then comes as a relief for this subgroup. Along similar lines, the job tasks of higher educated women might on average be more beneficial for maintaining a good mental and physical condition, which could make retiring relatively less beneficial.

Our results are line with, for example, Blake and Garrouste (2017), who find detrimental effects of a French pension reform on health. They find strong negative effects on self-reported health for low educated individuals only. Exploiting cross-country variation in retirement rules, Mazzonna and Peracchi (2017) find evidence of an immediate and positive effect of retirement on health if jobs are physically burdensome. Yet, in the mid- and long-run, they find negative effects of retirement. Also using SOEP data but a different identification strategy than our paper, Eibich (2015) finds beneficial effects of retirement with physically straining occupations benefiting most. Estimating the reform effects of an increase of the UK state pension age of women, Carrino, Glaser, and Avendano (2018) find a decrease in the physical and mental health of women with routine-manual jobs. Together with the effect heterogeneity found in this study, those results raise important questions about the impact of demographic change and pension reforms on inequality.

Nevertheless, there also exists a body of literature finding detrimental health effects of retirement. Even after taking into account the various health measures, the different time-scopes, and the diverse identification strategies, it is difficult to reconcile some parts of the results; see also the overview article of Eibich (2014).

4.5.3 Sensitivity analyses

In our baseline specification, the second stage equation is estimated on a combined survey data set of SHARE and SOEP. As a robustness check, the first row of Table 4.7 shows the estimated second stage effect of retirement on health based on the SOEP sample only. The causal effect of retirement on health differs slightly from the effect reported in Table 4.5 estimated from the joint survey data of SOEP and SHARE. Dropping the SHARE sample in the second stage increases the beneficial effect by 0.2, a quarter of the original effect size. The effect on the full SOEP sample is now significant on the 10% level. For the low education group, we find very similar results compared to our main specification, as can be seen from comparison of the third row of Table 4.7 with Table 4.6.

As pointed out in Section 4.3, several German states enacted school reforms differentially affecting cohorts born during the 1950s. To analyze the extent to which our results are confounded by these educational reforms, we compute an additional school reform control variable. This dummy variable indicates whether the individual was affected by the expansion of compulsory schooling in her state.[13] The state of residence is not part of the SHARE data set. Therefore, we run the second stage for this robustness check on the SOEP sample only. Comparing the first and the second row of Table 4.7, additionally adding the school reform dummy to the set of covariates in the first and second stage equations has virtually no effect on the coefficient of interest. The same exercise is repeated for the low educated in the third and fourth rows of Table 4.7. Additionally adding the school reform dummy only causes a minor change in the effect. We conclude that the expansion of compulsory schooling does not bias the estimates of our main specification.

In a further robustness check, we examine the sensitivity of our results to the use of the TS2SLS method. We run the entire estimation of both stages on the survey sample (SOEP and SHARE) as a standard 2SLS estimation and do not use the administrative records of the pension insurance in the first stage. The results of this exercise are shown in the lower panel of Table 4.7. The first stage effect of the reform on the retirement status slightly increases to about 15.6 percentage points in the full sample and to about 17.4 percentage points in the sample of less educated women; see the right section of rows 1 and 3 in the lower panel compared to the first stage effects as reported in Table 4.4. These effects are estimated with great precision, but standard errors are still four times larger compared with the first stage estimates based on the administrative data. Consequently, the second stage point estimates of the average effect of retirement on health barely change but precision decreases; see second stage coefficients in rows 1 and 3 of the lower panel of Table 4.7 in comparison to the coefficients reported in Tables 4.5 and 4.6. Point estimates from a standard 2SLS using survey data only are consistent with the findings of our main specification, which underlines the robustness of our main conclusions.

In an additional specification, we add an indicator for observations from the SOEP (vs. the SHARE sample) to control for systematic differences in survey design and questionnaires (see the discussion in Section 4.4). Comparing rows 1 and 2 in the lower panel of Table 4.7 shows that the full sample estimates are practically unaffected by the inclusion of a control variable that indicates the source of the survey data. Comparing rows 3 and 4, the same applies to the estimates of the less educated. This is despite the survey indicator being significant in the second stage regression (not shown).

[13]We proxy whether an individual was affected by the school reform by assuming current residence corresponds to the state in which the individual went to school.

Table 4.7: Second stage results, robustness checks

	2^{nd} Stage					1^{st} Stage		
	Coeff.	p-Value	[90% CI]		N	Coeff.	p-Value	N
2^{nd} stage estimated on SOEP data only								
Full Sample	-1.033	0.075	-1.988	-0.077	2,361	-0.153	0.00000	607,104
Full Sample, School reform	-1.039	0.068	-1.974	-0.104	2,361	-0.153	0.00000	607,104
Low Educ.	-1.279	0.093	-2.530	-0.027	1,645	-0.160	0.00000	487,404
Low Educ., School reform	-1.264	0.087	-2.481	-0.048	1,645	-0.162	0.00000	487,404
1^{st} (and 2^{nd}) stage on survey data								
Full Sample	-0.797	0.191	-1.800	0.206	2,894	-0.156	0.00015	2,894
Full Sample, Survey Dummy	-0.841	0.162	-1.831	0.149	2,894	-0.158	0.00017	2,894
Low Educ.	-1.153	0.133	-2.416	0.110	2,047	-0.174	0.00010	2,047
Low Educ., Survey Dummy	-1.144	0.135	-2.404	0.116	2,047	-0.174	0.00013	2,047

Note. Women Age 60-62 of cohorts 1950-1953. SOEP 2010 to 2016, SHARE waves 4 to 6. VSKT2016 (FDZ-RV), years 2010 to 2016. Standard errors clustered on month of birth.

Differences between SOEP and SHARE are present but not systematic and not confounding our design.

In our main specification, we treat the outcome variable, surveyed on a five point scale, as continuous. In our last robustness check, we assess the sensitivity of our results to this assumption. Thus, we define two alternative outcome measures as binary health variables indicating good health.[14] The first is defined as health rated as either 1 or 2 on the original scale. The second variable additionally includes health being rated as a 3. The results of this exercise are shown in Appendix Table 4.A.9 for the full sample and Table 4.A.10 for women without higher education. The results show qualitatively a very similar pattern as our main specification. The effect of retirement on health is estimated to be beneficial, with large but insignificant point estimates, in the full sample. In the lower education subsample, the point estimates turn larger and for the first indicator weakly significant, suggesting effect heterogeneity.

4.6 Conclusion

This paper contributes to the literature on the effects of retirement on health by exploiting exogenous variation from a large German pension reform that led to an increase in the ERA for women from age 60 to 63. In the empirical analysis, we use the strengths of both precise administrative data from the German pension insurance and detailed survey data from SOEP and SHARE. We estimate a TS2SLS model using administrative data in the first stage and survey data in the second stage.

As expected, and in line with previous studies, the 1999 pension reform induces a strong first stage effect on retirement. The second stage effect of retirement on self-reported health is non-detrimental. In fact, we find suggestive evidence that retirement is even beneficial. Yet, confidence intervals are large and a cautious interpretation can only rule out economically significant detrimental effects of retirement on health. Our results suggest effect heterogeneity. Point estimates of low educated women consistently indicate a more beneficial effect of retirement on health compared to the full sample point estimates. Further, estimates of low educated women are also more often statistically significant or are significant at a higher level of confidence.

[14]Note, these binary measures are defined as indicating good health, whereas the original health measure on the five point scale is defined on a reverse scale (higher values correspond to worse health status).

The findings of our study are robust to a range of sensitivity analyses. Most importantly, the effects neither rely on the two-sample method, nor are they biased by the compulsory school reforms that affected some of the cohorts in our sample.

Despite the contribution we make, the size of the survey data and the limited information available in the administrative data set some limitations in terms of possible heterogeneity analyses with respect to occupations, household characteristics, and further dimensions. Finally, related to the existing literature that distinguishes between short- and long-run effects of retirement, an extension of our analysis in a few years could be insightful.

Appendix

Monte Carlo simulation: Cluster robust standard errors

In a Monte Carlo study, we test whether the heteroskedasticity-robust variance estimator derived by Pacini and Windmeijer (2016) can be readily transformed into a cluster-robust version. To this end, we exchange the heteroskedasticity-robust variance estimators of first stage and reduced form inside the general TS2SLS variance estimator formula of Pacini and Windmeijer by their cluster-robust counterparts. Using 10,000 replications, our adaption of the analytical variance estimator and a bootstrapping method produces virtually identical standard errors. In the following, we describe the data generating process of our Monte Carlo exercise, briefly discuss the TS2SLS variance estimator as proposed by Pacini and Windmeijer (2016), describe our changes, and discuss the results of our simulation.

With three instruments, two unobserved/endogenous explanatory variables, and another exogenous explanatory variable, the model of the Monte Carlo study is more complex than the model used in the empirical part of this paper. Therefore, the simulation study applies to a more general context than our particular empirical application of the TS2SLS estimator. Further, our Monte Carlo setup generally follows the systematics, flexibility, and notation of the Monte Carlo simulation of Pacini and Windmeijer (2016).

Assume we were to estimate the following equation,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 w + \epsilon, \tag{4.A.3}$$

but x_1 and x_2 are not observed in the same data set as y. Yet, x_1 and x_2 can be instrumented: y is observed together with instruments z_1, z_2, z_3, and exogenous variable w in data set 1. Further, x_1 and x_2 are observed together with variables z_1, z_2, z_3, and w in data set 2. To ease notation, we define $X = (\mathbf{1}, x_1, x_2, w)$, $Z = (\mathbf{1}, z_1, z_2, z_3, w)$. Z is observed in data set 1 and data set 2, but observations are not identical. In fact, observations are distinct and independent but come from the same data generating process. Therefore, we use Z_1 and Z_2 for clarity if the data set matters. We estimate equations (4.A.4) and (4.A.5) on data set 2 to eventually predict the unobserved x_1 and x_2 in data set 1. The predictions later will be denoted as $\widehat{X}_1 = (\mathbf{1}, \widehat{x}_1, \widehat{x}_2, \widehat{w})$. Because exogenous w is observed in both data sets, the fitted values of w in $\widehat{X}_1$ will be w itself.

$$x_1 = \gamma_0 + \gamma_1 z_1 + \gamma_2 z_2 + \gamma_3 z_3 + \gamma_4 w + u_1 \tag{4.A.4}$$

$$x_2 = \gamma_5 - \gamma_6 z_1 + \gamma_7 z_2 + \gamma_8 z_3 + \gamma_9 w + u_2 \tag{4.A.5}$$

In the data generating process, the coefficient vector $\boldsymbol{\beta}$ is set to $(.9, -.6)'$, and $\boldsymbol{\gamma}$ is set to

$$\boldsymbol{\gamma} = (0, -.6, -1.6, 1.6, .2, 0, .5, -.5, -1.8, .4)'.$$

$Z = (\mathbf{1}, z_1, z_2, z_3, w)$ is drawn from a uniform distribution ranging from 0 to 1. The error structure exhibits heteroskedasticity and correlation within clusters c. u_1, u_2, and ϵ feature a heteroskedastic individual error component and a cluster-specific error component that, again, exhibits individual heteroskedasticity,

$$\begin{aligned} u_{1,ic} =& \nu_{1,i}\sqrt{exp(\alpha_1 z_{1i} + \alpha_2 z_{2i} + \alpha_3 z_{3i})} \\ &+ \mu_{1,c}\sqrt{exp(\alpha_4 z_{1i} + \alpha_5 z_{2i} + \alpha_6 z_{3i})}, \qquad (4.A.6) \\ u_{2,ic} =& \nu_{2,i}\sqrt{exp(\alpha_1 z_{1i} + \alpha_2 z_{2i} + \alpha_3 z_{3i})} \\ &+ \mu_{2,c}\sqrt{exp(\alpha_4 z_{1i} + \alpha_5 z_{2i} + \alpha_6 z_{3i})}, \qquad (4.A.7) \\ \epsilon_{ic} =& \nu_{3,i}\sqrt{exp(\alpha_1 z_{1i} + \alpha_2 z_{2i} + \alpha_3 z_{3i})} \\ &+ \mu_{3,c}\sqrt{exp(\alpha_4 z_{1i} + \alpha_5 z_{2i} + \alpha_6 z_{3i})}. \qquad (4.A.8) \end{aligned}$$

Heteroskedasticity is introduced through $\boldsymbol{\alpha} = (1, -1, .5, 1, .5, 1)'$. $\boldsymbol{\nu}$ is individual-specific, whereas $\boldsymbol{\mu}$ introduces within-cluster correlation; that is, $\boldsymbol{\nu}$ is also found in the Monte Carlo simulation of Pacini and Windmeijer, whereas $\boldsymbol{\nu}$ is only found in our simulation. Both $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ are drawn from multivariate normal distributions with parameters $\rho_1 = .3$ and $\rho_2 = -.2$,

$$\boldsymbol{\nu} = \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1\rho_2 \\ \rho_2 & \rho_1\rho_2 & 1 \end{pmatrix} \right), \qquad (4.A.9)$$

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1\rho_2 \\ \rho_2 & \rho_1\rho_2 & 1 \end{pmatrix} \right). \qquad (4.A.10)$$

After generating data based on the above specified relations, we split our sample in two parts, $N_1 = 800$ in 80 Clusters and $N_2 = 1200$ in 120 Clusters. Then, we delete variables x_1 and x_2 in data set 1. Further, we delete y in data set 2. Eventually, we estimate β_0, β_1, β_2 and β_3 as $\widehat{\beta}_{ts2sls}$ in a two-step procedure using the two distinct data sets and OLS. As derived in Pacini and Windmeijer (2016), the TS2SLS estimator of β is given by

$$\begin{aligned} \widehat{\beta}_{ts2sls} &= (\widehat{X}'\widehat{X}')^{-1}\widehat{X}_1' y_1 = (\widehat{\Pi}' Z_1' Z_1 \widehat{\Pi}')^{-1}\widehat{\Pi}' Z_1' y_1 \\ &= (\widehat{\Pi}' Z_1' Z_1 \widehat{\Pi}')^{-1}\widehat{\Pi}' Z_1' Z_1 \widehat{\pi}, \qquad (4.A.11) \end{aligned}$$

with $\widehat{\pi}$ being the estimated coefficient vector of a regression of y on Z_1. Correspondingly, $\widehat{\Pi}$ is the $k_z \times 3$ matrix of coefficient estimates from equations (4.A.4) and (4.A.5) used to predict X in data set 1 as $\widehat{X}_1$. k_z equals the number of included and excluded instruments, that is, z_1, z_2, z_3, w, and a constant. The first column of $\widehat{\Pi}$ contains estimates of γ_0 to γ_4, the second column contains estimates of γ_5 to γ_9, and the third column is defined as $(0, 0, 0, 0, 1)'$ because w is also observed in data set 1 and the fitted values of w in $\widehat{X}_1$ will be w itself. The corresponding variance estimator is computed according to Eq. (4.A.12) as described in Pacini and Windmeijer (2016),

$$\widehat{Var}(\widehat{\beta}_{ts2sls}) = \widehat{C}\widehat{Var}(\widehat{\pi})\widehat{C}' + (\widehat{\beta}'_{ts2sls} \otimes \widehat{C}) \times \widehat{Var}(\widehat{\Pi})(\widehat{\beta}_{ts2sls} \otimes \widehat{C}') \tag{4.A.12}$$

$$\text{with } \widehat{C} = (\widehat{X}'\widehat{X}')^{-1}\widehat{X}'Z_1.$$

Using robust specifications of $\widehat{Var}(\widehat{\pi})$ and $\widehat{Var}(\widehat{\Pi})$ results in a heteroskedasticity-robust variance estimator for $\widehat{Var}(\widehat{\beta}_{ts2sls})$ as shown by Pacini and Windmeijer. In contrast, we use cluster-robust versions of $\widehat{Var}(\widehat{\pi})$ and $\widehat{Var}(\widehat{\Pi})$[15] to gain a cluster-robust variance estimator $\widehat{Var}(\widehat{\beta}_{ts2sls})$ in Eq. (4.A.12).

Our simulation exercise shows that standard errors from the analytical estimator and a bootstrapping method produce similar results. When averaging the estimated cluster-robust standard errors computed from the analytical expression in Eq. (4.A.12) over 10,000 independent sets of data, the mean is close to the empirical standard deviation of the 10,000 estimates of β_1 and β_2; see Table 4.A.8. Put differently, in the presence of cluster-specific autocorrelation in error terms, inference based on the analytical expression of the variance estimator produces results comparable to a simple bootstrap routine. The mean standard errors computed from the robust variance estimator as described in Pacini and Windmeijer (2016) and the non-robust estimator are included for comparison. On average, these other estimators result in lower standard errors.

To illustrate how the use of a wrong variance estimator can lead to an inflation of false positives, we use Wald tests to reject the null hypothesis that the estimated coefficients equal the true coefficient values, that is, $H_0 : \widehat{\beta}_1 = 0.9$ and $H_0 : \widehat{\beta}_2 = -0.6$. Using standard errors computed from the cluster-robust variance estimator leads to a rejection of the null hypothesis at the 5% level of confidence in 5.0% and 4.7% of the simulations; see the last column of Table 4.A.8. Using standard and heteroskedasticity-robust

[15]In Stata, the cluster-robust variance estimators of $\widehat{\pi}$ and $\widehat{\Pi}$ can be produced from the gmm or the `suest` command.

Table 4.A.8: Monte Carlo results, variance estimators

	Mean $\widehat{\beta}$	Std. Dev.	Mean Std. Err.			Wald Rejections		
			Hom.	Rob.	Cluster	Hom.	Rob.	Cluster
β_1	0.893	0.149	0.107	0.126	0.145	0.153	0.087	0.050
β_2	-0.601	0.220	0.131	0.167	0.214	0.234	0.126	0.047

Note. 10,000 Simulations, data set 1 with $N_1 = 800$ in 80 Clusters, data set 2 with $N_2 = 1,200$ in 120 Clusters. Mean $\widehat{\beta}$ from TS2SLS. *Mean Std. Err.* denotes the mean std. error (from the standard, the robust and the cluster-robust variance estimator). Under *Wald Rejections*, the share of rejections of the null hypothesis that the estimated coefficient equals its true value are reported, $H_0 : \widehat{\beta_k} = \beta_k$, 5% level of confidence (t=1.963).

standard errors to compute the Wald test-statistics leads to over-rejection, see the third to last and second to last column of Table 4.A.8.

The Stata code of the simulation study, including the code for the TS2SLS coefficient and variance estimator, can be accessed under github.com/setgeton/cluster-se, parts of it are based on the online appendix of Pacini and Windmeijer (2016).

Other results

Table 4.A.9: Second stage results, full sample, binary health indicators

	Health<=2	p-Value	Health<=3	p-Value
Retirement	0.357	0.123	0.402	0.165
Birth Cohort	0.039	0.166	0.032	0.133
Birth Cohort X Treated	-0.024	0.494	-0.021	0.546
East Germany	-0.052	0.291	-0.032	0.552
Age in Months	-0.002	0.077	-0.003	0.006
Constant	1.766	0.027	2.701	0
N	2894		2894	

Note. Women Age 60-62 of cohorts 1950-1953 from SOEP 2010 to 2016 and SHARE waves 4 to 6 data. VSKT2016 (FDZ-RV), years 2010 to 2016. Standard errors clustered on month of birth.

Table 4.A.10: Second stage results, low education, binary health indicators

	Health<=2	p-Value	Health<=3	p-Value
Retirement	0.518	0.077	0.456	0.172
Birth Cohort	0.027	0.371	0.026	0.285
Birth Cohort X Treated	-0.013	0.757	-0.012	0.780
East Germany	-0.087	0.160	-0.054	0.423
Age in Months	-0.003	0.035	-0.003	0.020
Constant	2.322	0.013	2.583	0.001
N	2047		2047	

Note. Women Age 60-62 of cohorts 1950-1953 from SOEP 2010 to 2016 and SHARE waves 4 to 6 data. VSKT2016 (FDZ-RV), years 2010 to 2016. Standard errors clustered on month of birth.

5 Conclusion

This dissertation comprises three self-contained research articles, each examining different aspects of pension reforms. In its entirety, the dissertation sheds some light on the evolution of inequality in times of generosity decreasing pension reforms.

In the first study, I use a structural model to *ex-ante* simulate employment outcomes of an increase of the normal retirement age (NRA) in Germany from 65 to 67. In the baseline scenario, age at employment exit and retirement age increase by only 0.6 years. Resulting pension benefits decline by 2.0%. Widespread reform effectiveness is hampered by the heterogeneous availability of jobs. The risk of involuntary job loss varies along socio-demographic characteristics. In consequence, old-age income inequality increases.

In the second study, I present theoretical and empirical evidence that an increase of the early retirement age (ERA) not necessarily increases savings. The theoretical section shows that the effect of a decline of pension generosity on savings rates is ambiguous. The effect sign crucially depends on the corresponding employment effect. In the empirical section, I estimate the effect of an increase of the ERA using detailed households savings information and variation from a strong and not gradually phased-in increase of the ERA in Germany. An increase of the ERA does not lead to increased savings rates of working age individuals. Findings are in line with the anticipation of prolonged careers. Further, results indicate that low wealth and high education households show more negative reform effects on savings rates. As potential mechanism for the more negative point estimate of high educated households, I discuss better chances of continued employment. Further, financial literacy could allow to more easily re-optimize savings plans. The more negative point estimate of low wealth households is potentially caused by a high marginal utility of consumption and a relatively low disutility of prolonged employment.

Lastly, the causal effect of retirement on health is estimated. As in the previous study, identification stems from the exogenous and discontinuous increase of the ERA of German women. In a two-sample two-stage least squares estimation framework, precise administrative data is combined with survey data on self-reported health. Results show no signs of a detrimental effect of retirement on health. In fact, results suggest that especially less educated individuals benefit from retirement. Pension reforms that restrict access to retirement are, therefore, prone to disproportionately harm less educated individuals.

The separate results of this dissertation give a multi-dimensional view of the effects of pension reforms. Throughout the dissertation I find effect heterogeneity or suggestive evidence of effect heterogeneity along socio-demographic characteristics. In particular, the effect of the increase of the NRA on income inequality, the heterogeneous effect of an increase of the ERA on savings rates, and the effect of retirement on health casts serious doubts on whether pension reforms aiming at financial viability of the pension system are necessarily also socially balanced. A varying ability to cope with pension reforms could exacerbate societal imbalances. This notion is worrisome, because income inequality in the developed world is already high (Cingano 2014), the effect of income inequality on population health and wellbeing is robust (Pickett and Wilkinson 2015), and pension reforms in many OECD countries are still ongoing or planned for the near future (OECD 2017).

Therefore, I end this dissertation with two short policy recommendations. Firstly, and as a more general remark, I recommend that distributional aspects play a crucial role in the design of pension reforms. In particular, regulatory impact analysis (*Gesetzesfolgenabschätzung*) should, next to financial costs of a law, always contain explicit statements on the expected distributional effects and name adversely affected socio-demographic groups. Thereby, the quality and adequacy of the public debate about particular reforms could be improved at an early stage. Furthermore, policy makers should enshrine adequate financial means and general procedures for practicable and routine policy evaluation in the law of reforms to ensure effective evaluation of the actual effects of reforms. Second, policy makers should try to protect vulnerable groups from potentially detrimental effects of pension reforms through carefully targeted regulations and initiatives while accounting for moral hazard. In particular, I recommend efforts to increase the old-age employability of disadvantaged groups, efforts to increase participation of disadvantaged groups in subsidized savings, efforts to improve working conditions and rehabilitation programs, and, lastly, efforts to strengthen retirement options of elderly individuals with delicate health.

Bibliography

Alessie, Rob, Viola Angelini, and Peter van Santen (2013). “Pension wealth and household savings in Europe: Evidence from SHARELIFE”. In: *European Economic Review* 63, pp. 308 –328.

Alik-Lagrange, Arthur and Tobias Schmidt (2015). *The pattern of home ownership across cohorts and its impact on the net wealth distribution: Empirical evidence from Germany and the US.* Discussion Papers 11/2015. Deutsche Bundesbank, Research Centre.

Angrist, Joshua D. and Alan B. Krueger (1992). “The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples”. In: *Journal of the American Statistical Association* 87.418, pp. 328–336.

Arellano, Manuel and Costas Meghir (1992). “Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets”. In: *The Review of Economic Studies* 59.3, pp. 537–559.

Atalay, Kadir and Garry F. Barrett (2015). “The Impact of Age Pension Eligibility Age on Retirement and Program Dependence: Evidence from an Australian Experiment”. In: *The Review of Economics and Statistics* 97.1, pp. 71–87.

Attanasio, Orazio P. and Agar Brugiavini (2003). “Social Security and Households’ Saving”. In: *The Quarterly Journal of Economics* 118.3, pp. 1075–1119.

Attanasio, Orazio P. and Susann Rohwedder (2003). “Pension Wealth and Household Saving: Evidence from Pension Reforms in the United Kingdom”. In: *American Economic Review* 93.5, pp. 1499–1521.

Autor, David H (2014). “Skills, education, and the rise of earnings inequality among the "other 99 percent"”. In: *Science (New York, N.Y.)* 344.6186, 843—851.

Behaghel, Luc and David M. Blau (2012). “Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age”. In: *American Economic Journal: Economic Policy* 4.4, pp. 41–67.

Bernheim, B. Douglas, Jonathan Skinner, and Steven Weinberg (2001). “What Accounts for the Variation in Retirement Wealth among U.S. Households?” In: *American Economic Review* 91.4, pp. 832–857.

Blake, Helene and Clementine Garrouste (2017). *Collateral effects of a pension reform in France.* Health, Econometrics and Data Group (HEDG) Working Papers 12/16. Department of Economics, University of York.

Blau, David M. (2008). “Retirement and Consumption in a Life Cycle Model”. In: *Journal of Labor Economics* 26, pp. 35–71.

Bloemen, Hans, Stefan Hochguertel, and Jochem Zweerink (2013). *The causal effect of retirement on mortality: Evidence from targeted incentives to retire early.* Discussion Paper DP 08/2013-024. Netspar.

Blundell, Richard, John Ham, and Costas Meghir (1987). “Unemployment and Female Labour Supply”. In: *Economic Journal* 97.388a, pp. 44–64.

Bottazzi, Renata, Tullio Jappelli, and Mario Padula (2006). “Retirement expectations, pension reforms, and their impact on private wealth accumulation”. In: *Journal of Public Economics* 90.12, pp. 2187–2212.

Brown, Kristine M. (2013). “The link between pensions and retirement timing: Lessons from California teachers”. In: *Journal of Public Economics* 98, pp. 1 –14.

Bundesbank (2016). “Household wealth and finances in Germany: results of the 2014 survey”. In: *Monthly Report March 2016.* Vol. 68. No 3. Deutsche Bundesbank.

Burda, Michael C and Antje Mertens (2001). “Estimating wage losses of displaced workers in Germany”. In: *Labour Economics* 8.1, pp. 15 –41.

Bönke, Timm, Giacomo Corneo, and Holger Lüthen (2015). "Lifetime Earnings Inequality in Germany". In: *Journal of Labor Economics* 33.1, pp. 171–208.

Bönke, Timm, Daniel Kemptner, and Holger Lüthen (2018). "Effectiveness of early retirement disincentives: Individual welfare, distributional and fiscal implications". In: *Labour Economics* 51, pp. 25 –37.

Bönke, Timm, Markus Grabka, Carsten Schröder, Edward N. Wolff, and Lennard Zyska (2016). *The Joint Distribution of Net Worth and Pension Wealth in Germany*. SOEPpapers on Multidisciplinary Panel Data Research 853. DIW Berlin, The German Socio-Economic Panel (SOEP).

Börsch-Supan, Axel (2017a). "Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4". In: *Data set* Release version: 6.0.0. SHARE-ERIC, DOI: 10.6103/SHARE.w4.600.

Börsch-Supan, Axel (2017b). "Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5". In: *Data set* Release version: 6.0.0. SHARE-ERIC, DOI: 10.6103/SHARE.w5.600.

Börsch-Supan, Axel (2017c). "Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6". In: *Data set* Release version: 6.0.0. SHARE-ERIC, DOI: 10.6103/SHARE.w6.600.

Börsch-Supan, Axel and Christina B. Wilke (2004). *The German Public Pension System: How it was, how it will be.* Working paper series 10525. NBER.

Börsch-Supan, Axel, Reinhold Schnabel, Simone Kohnz, and Giovanni Mastrobuoni (2004). "Micro-modeling of retirement decisions in Germany". In: *Social security programs and retirement around the world: Micro-estimation.* University of Chicago Press, pp. 285–344.

Börsch-Supan, Axel, Martina Brandt, Christian Hunkler, Thorsten Kneip, Julie Korbmacher, Frederic Malter, Barbara Schaan, Stephanie Stuck, and Sabrina Zuber (2013). "Data resource profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)". In: *International Journal of Epidemiology* 42, pp. 992–1001.

Börsch-Supan, Axel, Tabea Bucher-Koenen, Michela Coppola, and Bettina Lamla (2015). "Savings in Times of Demographic Change: Lessons from the German Experience”. In: *Journal of Economic Surveys* 29.4, pp. 807–829.

Card, David, Jörg Heining, and Patrick Kline (2013). “Workplace Heterogeneity and the Rise of West German Wage Inequality”. In: *The Quarterly Journal of Economics* 128.3, pp. 967–1015.

Carrino, Ludovico, Karen Faria Glaser, and Mauricio Pabon Avendano (2018). *Later pension, poorer health?: Evidence from the New State Pension age in the UK*. Working Paper Series 87575. Harvard Center for Population and Development Studies.

Chan, Sewin and Ann Huff Stevens (2004). “How Does Job Loss Affect the Timing of Retirement?” In: *The B.E. Journal of Economic Analysis & Policy* 3.1, pp. 1–26.

Charles, Kerwin (2004). “Is retirement depressing?: Labor force inactivity and psychological well-being in later life”. In: *Research in Labor Economics*. Ed. by Solomon W. Polachek. Vol. 23: Accounting for worker well-being. Emerald Group Publishing Limited, pp. 269–299.

Chetty, Raj, John N. Friedman, Søren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen (2014). “Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark”. In: *The Quarterly Journal of Economics* 129.3, pp. 1141–1219.

Cingano, Federico (2014). “Trends in Income Inequality and its Impact on Economic Growth”. In: 163.

Coe, Norma B. and Gema Zamarro (2011). “Retirement effects on health in Europe”. In: *Journal of Health Economics* 30.1, pp. 77–86.

Coleman, David (2008). “The demographic effects of international migration in Europe”. In: *Oxford Review of Economic Policy* 24.3, pp. 452–476.

De Grip, Andries, Maarten Lindeboom, and Raymond Montizaan (2011). “Shattered dreams: The effects of changing the pension system late in the game”. In: *The Economic Journal* 122, pp. 1–25.

De Nardi, Mariacristina, Eric French, and John Bailey Jones (2016). "Savings after Retirement: A Survey". In: *Chicago Fed Letter*.

Deutsche Rentenversicherung Bund (2016). *Versicherungskontenstichprobe (VSKT) 2016*. Source: FDZ-RV, on-site access.

Duggan, Mark, Perry Singleton, and Jae Song (2007). "Aching to retire? The rise in the full retirement age and its impact on the social security disability rolls". In: *Journal of Public Economics* 91.7-8, pp. 1327–1350.

Dustmann, Christian, Bernd Fitzenberger, Uta Schönberg, and Alexandra Spitz-Oener (2014). "From Sick Man of Europe to Economic Superstar: Germany's Resurgent Economy". In: *Journal of Economic Perspectives* 28.1, pp. 167–88.

Ebbinghaus, Bernhard and Jonas Radl (2015). "Pushed out prematurely? Comparing objectively forced exits and subjective assessments of involuntary retirement across Europe". In: *Research in Social Stratification and Mobility* 41, pp. 115 –130.

ECB (04/2013). *The Eurosystem Household Finance and Consumption Survey - Results from the first wave*. Statistics Paper Series 2. European Central Bank.

Eibich, Peter (2014). *The Health Effects of Retirement*. DIW Roundup: Politik im Fokus 48. DIW Berlin, German Institute for Economic Research.

Eibich, Peter (2015). "Understanding the effect of retirement on health: Mechanisms and heterogeneity". In: *Journal of Health Economics* 43, pp. 1–12.

Engels, Barbara, Johannes Geyer, and Peter Haan (2017). "Pension incentives and early retirement". In: *Labour Economics* 47, pp. 216 –231.

Etgeton, Stefan (2018). "The effect of pension reforms on old-age income inequality". In: *Labour Economics* 53, pp. 146 –161.

Eurostat (03/2018). *Unemployment rates by sex, age and educational attainment level*. URL: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsq_urgaed\&lang=en.

Fehr, Hans, Manuel Kallweit, and Fabian Kindermann (07/2012). “Pension reform with variable retirement age: a simulation analysis for Germany”. In: *Journal of Pension Economics and Finance* 11 (03), pp. 389–417.

Feldstein, Martin (1974). “Social Security, Induced Retirement, and Aggregate Capital Accumulation”. In: *Journal of Political Economy* 82.5, pp. 905–926.

Feldstein, Martin (1976). “Social Security and Saving: The Extended Life Cycle Theory”. In: *The American Economic Review* 66.2, pp. 77–86.

Feng, Jin, Lixin He, and Hiroshi Sato (2011). “Public pension and household saving: Evidence from urban China”. In: *Journal of Comparative Economics* 39.4, pp. 470 –485.

Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter (2006). “Imputation Rules to Improve the Education Variable in the IAB Employment Subsample”. In: *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften* 126.3, pp. 405–436.

Fitzpatrick, Maria D. and Timothy J. Moore (2016). *The mortality effects of retirement: Evidence from social security eligibility at age 62.* Working Paper 2016-7. Center for Retirement Research at Boston College.

Frejka, Tomas and Tomáš Sobotka (2008). “Fertility in Europe: Diverse, delayed and below replacement”. In: *Demographic Research* Special Collection 7: Childbearing Trends and Policies in Europe.19(3), pp. 15–46.

French, Eric (2005). “The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour”. In: *Review of Economic Studies* 72.2, pp. 395–427.

Gale, William G. (1998). “The Effects of Pensions on Household Wealth: A Reevaluation of Theory and Evidence”. In: *Journal of Political Economy* 106.4, pp. 706–723.

Geyer, Johannes and Clara Welteke (2017). *Closing Routes to Retirement: How Do People Respond?* Discussion Papers of DIW Berlin 1653. DIW Berlin, German Institute for Economic Research.

Geyer, Johannes, Peter Haan, Anna Hammerschmid, and Michael Peters (2018). *Labor Market and Distributional Effects of an Increase in the Retirement Age*. Rationality and Competition Discussion Paper Series 101. CRC TRR 190 Rationality and Competition.

Godard, Mathilde (2016). “Gaining weight through retirement? Results from the SHARE survey”. In: *Journal of Health Economics* 45, pp. 27–46.

Goebel, Jan, Markus M Grabka, Stefan Liebig, Martin Kroh, David Richter, Carsten Schröder, and Jürgen Schupp (2018). “The German Socio-Economic Panel (SOEP)”. In: *Jahrbücher für Nationalökonomie und Statistik*.

Gourinchas, Pierre-Olivier and Jonathan A. Parker (2002). “Consumption over the Life Cycle”. In: *Econometrica* 70.1, pp. 47–89.

Grogger, J. and C. Wunsch (2012). “Unemployment Insurance and Departures from Employment: Evidence from a German Reform”. In: *Mimeo, VU Amsterdam*.

Gustman, Alan L. and Thomas L. Steinmeier (2000). “Retirement in Dual-Career Families: A Structural Model”. In: *Journal of Labor Economics* 18.3, pp. 503–545.

Gustman, Alan L. and Thomas L. Steinmeier (2015). “Effects of social security policies on benefit claiming, retirement and saving”. In: *Journal of Public Economics* 129.(C), pp. 51 –62.

Haan, Peter, Daniel Kemptner, and Holger Lüthen (10/2017). *The Rising Longevity Gap by Lifetime Earnings: Distributional Implications for the Pension System*. IZA Discussion Papers 11121. Institute for the Study of Labor (IZA).

Haan, Peter and Victoria Prowse (2014). “Longevity, life-cycle behavior and pension reform”. In: *Journal of Econometrics* 178, Part 3, pp. 582 –601.

Haider, Steven J. and Melvin Stephens (2007). “Is There a Retirement-Consumption Puzzle? Evidence Using Subjective Retirement Expectations”. In: *The Review of Economics and Statistics* 89.2, pp. 247–264.

Hairault, Jean-Olivier, Thepthida Sopraseuth, and François Langot (2010). "Distance to Retirement and Older Workers' Employment: The Case for Delaying the Retirement Age". In: *Journal of the European Economic Association* 8.5, pp. 1034–1076.

Hanel, Barbara (2010). "Financial incentives to postpone retirement and further effects on employment - Evidence from a natural experiment". In: *Labour Economics* 17.3, pp. 474 –486.

Hanel, Barbara and Regina T. Riphahn (2012). "The timing of retirement — New evidence from Swiss female workers". In: *Labour Economics* 19.5, pp. 718 –728.

Heckman, J. and B. Singer (1984). "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data". In: *Econometrica* 52.2, pp. 271–320.

Hochfellner, Daniela, Dana Müller, and Anja Wurdack (2012). "Biographical data of social insurance agencies in Germany–improving the content of administrative data". In: *Schmollers Jahrbuch* 132, pp. 443–451.

Idler, Ellen L. and Yael Benyamini (1997). "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies". In: *Journal of Health and Social Behavior* 38.1, pp. 21–37.

Imbens, Guido W. and Thomas Lemieux (2008). "Regression discontinuity designs: A guide to practice". In: *Journal of Econometrics* 142, pp. 615–635.

Inoue, Atsushi and Gary Solon (2010). "Two-Sample Instrumental Variables Estimators". In: *The Review of Economics and Statistics* 92.3, pp. 557–561.

Keane, Michael P., Petra E. Todd, and Kenneth I. Wolpin (2011). "Chapter 4 - The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications". In: ed. by Orley Ashenfelter and David Card. Vol. 4, Part A. Handbook of Labor Economics. Elsevier, pp. 331 –461.

Kemptner, Daniel, Hendrik Jürges, and Steffen Reinhold (2011). "Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany". In: *Journal of Health Economics* 30, pp. 340–354.

Klevmarken, Anders (1982). *Missing Variables and Two-Stage Least-Squares Estimation from More than One Data Set.* IUI Working Paper 62. Stockholm: The Research Institute of Industrial Economics.

Kolodziej, Ingo and Pilar Garcia-Gomez (2017). *The causal effects of retirement on mental health: Looking beyond the mean effects.* Ruhr Economics Papers 668. RWI RUB TU Dortmund U Duisburg-Essen.

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song (2010). "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data Since 1937*". In: *The Quarterly Journal of Economics* 125.1, pp. 91–128.

Krueger, Alan B. and Jörn-Steffen Pischke (1992). "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation". In: *Journal of Labor Economics* 10.4, pp. 412–437.

Kuhn, Andreas, Stefan Staubli, Jean-Philippe Wuellrich, and Josef Zweimüller (2018). *Fatal Attraction? Extended Unemployment Benefits, Labor Force Exits, and Mortality.* IZA Discussion Papers 11851. Institute for the Study of Labor (IZA).

Lachowska, Marta and Michał Myck (2018). "The effect of public pension wealth on saving and expenditure". In: *American Economic Journal: Economic Policy* 10.3, pp. 284–308.

Lalive, Rafael and Pierpaolo Parrotta (2017). "How does pension eligibility affect labor supply in couples?" In: *Labour Economics* 46.Supplement C, pp. 177 –188.

Lalive, Rafael and Stefan Staubli (2015). *How Does Raising Women's Full Retirement Age Affect Labor Supply, Income, and Mortality?* Working Paper NB 14-09. NBER Retirement Research Center.

Lee, David S. and Thomas Lemieux (2010). "Regression Discontinuity Designs in Economics". In: *Journal of Economic Literature* 48.2, pp. 281–355.

Leimer, Birgit (2017). *No "Honeymoon Phase": Whose health benefits from retirement and when.* Research Unit "Interdisciplinary Public Policy" Discussion Paper Series 1718. Gutenberg School of Management and Economics.

Lindahl, Mikael (2005). “Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income”. In: *Journal of Human Resources* 40.1, pp. 144–168.

Lindeboom, Maarten and Raymond Montizaan (2018). *Pension Reforms, Retirement Expectations and Realization, and Private Savings.* Academic Series DP 02/2018-012. Netspar.

Lüthen, Holger (2016). “Rates of Return and Early Retirement Disincentives: Evidence from a German Pension Reform”. In: *German Economic Review* 17.2, pp. 206–233.

Malter, Frederic and Axel Börsch-Supan, eds. (2015). *SHARE Wave 5: Innovations and methodology.* Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Malter, Frederic and Axel Börsch-Supan, eds. (2017). *SHARE Wave 6: Panel innovations and collecting dried blood spots.* Munich: MEA, Max Planck Institute for Social Law and Social Polic.

Manoli, Day, Kathleen J. Mullen, and Mathis Wagner (2015). “Policy Variation, Labor Supply Elasticities, and a Structural Model of Retirement”. In: *Economic Inquiry* 53.4, pp. 1702–1717.

Manoli, Day and Andrea Weber (11/2016). “Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions”. In: *American Economic Journal: Economic Policy* 8 (4).

Mastrobuoni, Giovanni (2009). “Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities”. In: *Journal of Public Economics* 93.11, pp. 1224 –1233.

Mazzonna, Fabrizio and Franco Peracchi (2017). “Unhealthy Retirement?” In: *Journal of Human Resources* 52.1, pp. 128–151.

Merkurieva, Irina (2016). *Late Career Job Loss and the Decision to Retire.* Discussion Paper Series, Department of Economics 2016/06. Department of Economics, University of St. Andrews.

Miilunpalo, Seppo, Ilkka Vuori, Pekka Oja, Matti Pasanen, and Helka Urponen (1997). "Self-rated health status as a health measure: the predictive value of self-reported health status on the use of physician services and on mortality in the working-age population". In: *Journal of clinical epidemiology* 50.5, pp. 517–528.

Moffitt, Robert (1983). "An Economic Model of Welfare Stigma". In: *The American Economic Review* 73.5, pp. 1023–1035.

Mulsant, Benoit H, Mary Ganguli, and Eric C. Seaberg (2015). "The Relationship Between Self-Rated Health and Depressive Symptoms in an Epidemiological Sample of Community-Dwelling Older Adults". In: *Journal of the American Geriatrics Society* 45.8, pp. 954–958.

OECD (2015). *Pensions at a Glance 2015: OECD and G20 indicators.* OECD Publishing, Paris.

OECD (2017). *Pensions at a Glance 2017: OECD and G20 indicators.* OECD Publishing, Paris.

Oeppen, Jim and James W. Vaupel (2002). "Broken Limits to Life Expectancy". In: *Science* 296.5570, pp. 1029–1031.

Pacini, David and Frank Windmeijer (2016). "Robust inference for the Two-Sample 2SLS estimator". In: *Economics Letters* 146, pp. 50 –54.

Pickett, Kate E. and Richard G. Wilkinson (2015). "Income inequality and health: A causal review". In: *Social Science & Medicine* 128, pp. 316 –326.

Piketty, Thomas and Emmanuel Saez (2013). "Chapter 7 - Optimal Labor Income Taxation". In: *Handbook of Public Economics.* Ed. by Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez. Vol. 5. Handbook of Public Economics. Elsevier, pp. 391 –474.

Poterba, James, Steven Venti, and David A. Wise (10/2015). *What Determines End-of-Life Assets? A Retrospective View.* NBER Working Papers 21682. National Bureau of Economic Research, Inc.

Puhani, Patrick A. and Falko Tabbert (2016). "The effects of pension changes on age of first benefit receipt: Regression discontinuity evidence from repatriated ethnic Germans". In: *Labour Economics* 38, pp. 12 –23.

Queisser, Monika and Edward Whitehouse (12/2006). *Neutral or Fair?: Actuarial Concepts and Pension-System Design*. OECD Social, Employment and Migration Working Papers 40. OECD Publishing.

Radl, Jonas (2013). "Labour Market Exit and Social Stratification in Western Europe: The Effects of Social Class and Gender on the Timing of Retirement". In: *European Sociological Review* 29.3, pp. 654–668.

Rentenreformgesetz 1999 (1997). "RRG 1999". In: *Bundesgesetzblatt* 1.85, pp. 2998–3028.

Rohwedder, Susann and Robert J. Willis (2010). "Mental retirement". In: *Journal of Economic Perspectives* 24.1, pp. 119–138.

Rust, John and Christopher Phelan (1997). "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets". In: *Econometrica* 65.4, pp. 781–831.

Schmitz, Hendrik (2011). "Why are the unemployed in worse health? The causal effect of unemployment on health". In: *Labour Economics* 18.1, pp. 71–78.

Seibold, Arthur (2018). "Essays on behavioral responses to social insurance and taxation". Doctoral thesis. London School of Economics and Political Science.

Selin, Håkan (2014). "The rise in female employment and the role of tax incentives. An empirical analysis of the Swedish individual tax reform of 1971". In: *International Tax and Public Finance* 21.5, pp. 894–922.

Smith, Sarah (2006). "The Retirement-Consumption Puzzle and Involuntary Early Retirement: Evidence from the British Household Panel Survey*". In: *The Economic Journal* 116.510, pp. C130–C148.

SOEP (2018). "Socio-Economic Panel, Data for years 1984-2016". In: *Data set* Version 33.1, doi:10.5684/soep.v33.1.

Stancanelli, Elena and Arthur van Soest (2012). “Retirement and Home Production: A Regression Discontinuity Approach”. In: *American Economic Review* 102.3, pp. 600–605.

Statistische Ämter des Bundes und der Länder (08/2018). *Einkommens- und Verbrauchsstichprobe*. Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder. URL: http://www.forschungsdatenzentrum.de/bestand/evs/index.asp.

Staubli, Stefan and Josef Zweimüller (2013). “Does raising the early retirement age increase employment of older workers?” In: *Journal of Public Economics* 108, pp. 17 –32.

Stock, James H and David A Wise (1990). “Pensions, the Option Value of Work, and Retirement”. In: *Econometrica* 58.5, pp. 1151–80.

Taber, Christopher R. (2000). “Semiparametric identification and heterogeneity in discrete choice dynamic programming models”. In: *Journal of Econometrics* 96.2, pp. 201 –229.

Train, Kenneth (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Vestad, Ola Lotherington (2013). “Labour supply effects of early retirement provision”. In: *Labour Economics* 25.Supplement C. European Association of Labour Economists 24th Annual Conference, Bonn, Germany, 20-22 September 2012, pp. 98 –109.

Wise, David A. (1997). “Retirement against the demographic trend: More older people living longer, working less, and saving less?” In: *Demography* 34.1, pp. 83–95.

Zhao, Lingxiao and Gregory Burge (2017). “Housing Wealth, Property Taxes, and Labor Supply among the Elderly”. In: *Journal of Labor Economics* 35.1, pp. 227–263.

Summary

This dissertation analyzes the effect of pension reforms on employment, retirement age, old-age income, savings, and health. Throughout, attention is paid to the heterogeneity of effects and potential implications for inequality are highlighted. The dissertation consists of three self-contained research articles.

In the first study, I analyze the effects of an increase of the normal retirement age (NRA) on employment, realized retirement age, and, in consequence, old-age income inequality. Many OECD countries are raising the NRA, thereby, making early retirement more costly. Whereas such reforms incentivize individuals to work longer, labor market frictions might partly undermine intended behavioral responses. Employing administrative data of West German men, I estimate a dynamic discrete choice model of work, unemployment and retirement allowing for labor market frictions. Involuntary job losses constrain individual choice sets to differing degrees along sociodemographic characteristics. A policy-simulation suggests that the behavioral response to an increase in the NRA from 65 to 67 is moderate, with an average delay of employment exits of only 0.6 years. Widespread reform effectiveness is hampered by the heterogeneous availability of jobs. Concerning the resulting pension benefits, poverty-vulnerable groups are hit hardest: Individuals with low education and blue collar employees suffer disproportionally. Old-age income inequality increases.

In the second study, I estimate the effects of an increase of the early retirement age (ERA) on savings rates. In theory, a decrease of pension generosity can have either a positive or negative effect on the savings rate. The sign of the effect depends on corresponding employment effects of the pension reform. In 1999, a reform lifted the ERA of German women born after 1951 by 3 years creating a discontinuity along birth cohorts. The reform discontinuity is used in an RDD setting to estimate the isolated effect of the ERA on savings rates. Estimation is based on detailed income and consumption data from a household survey. In contrast to the previous literature, an increase of the ERA is estimated to have a non-positive effect

on the savings rate. A non-positive effect is in line with the substantial extension of employment of affected cohorts. In particular, low wealth and highly educated households reduce savings rates upon being affected by the increase of the ERA. Potential mechanisms include a heterogeneous marginal utility of consumption and a lack of financial education of low educated households.

In the third study, the effects of retirement on health are analyzed. Again, the discontinuous increase in the ERA of German women is exploited in an RDD setting. The analysis is based on a two-sample two-stage regression framework using micro data of two well-established surveys and administrative records of the pension fund. The effect of retirement on self-reported health is examined paying particular attention to the effect heterogeneity across educational groups. Cautiously interpreting the estimates, we find a non-detrimental effect of retirement on health. For low educated women, estimates indicate a beneficial effect of retirement on health.

The three studies show that effects of pension reforms are heterogeneous. In terms of employment and retirement age, poverty-vulnerable groups react less to an increase of the NRA. In consequence, old-age income inequality increases. The analysis of the effect of the ERA suggests problems of low educated individuals to adjust savings. Further, low educated individuals show beneficial effects of retirement on health. Therefore, I conclude that a varying ability to cope with pension reforms might lead to a worsening of societal imbalances.

Zusammenfassung

Diese Dissertation analysiert die Auswirkungen von Rentenreformen auf Beschäftigung, Rentenalter, Alterseinkommen, Sparverhalten und Gesundheit. Dabei wird der Heterogenität der Effekte besondere Aufmerksamkeit geschenkt und auf mögliche Folgen für die Ungleichheit hingewiesen. Die Dissertation besteht aus drei eigenständigen Forschungsarbeiten.

In der ersten Studie werden die Auswirkungen einer Erhöhung des Regelrentenalters auf die Beschäftigung, das tatsächliche Rentenalter und, in der Konsequenz, die Einkommensungleichheit im Alter analysiert. Viele OECD-Länder erhöhen das Regelrentenalter, wodurch der Preis eines vorzeitigen Ruhestands steigt. Während solche Reformen den Einzelnen dazu anregen, länger zu arbeiten, könnten Arbeitsmarktfriktionen die beabsichtigten Verhaltensreaktionen teilweise untergraben. Unter Verwendung administrativer Daten westdeutscher Männer schätze ich ein dynamisches, discrete choice model von Beschäftigung, Arbeitslosigkeit und Ruhestand, das Friktionen auf dem Arbeitsmarkt berücksichtigt. Unfreiwillige Arbeitsplatzverluste schränken individuelle Auswahlmöglichkeiten entlang soziodemografischer Merkmale in unterschiedlichem Maße ein. Eine Politiksimulation deutet darauf hin, dass die Verhaltensreaktion auf einen Anstieg des Regelrentenalters von 65 auf 67 Jahre moderat ist: Das Ausscheiden aus der Beschäftigung verzögert sich im Durchschnitt um lediglich 0,6 Jahre. Eine breite Wirksamkeit der Reform wird durch die heterogene Verfügbarkeit von Arbeitsplätzen behindert. Bei den daraus resultierenden Rentenzahlungen sind armutsgefährdete Gruppen am stärksten betroffen: Menschen mit geringer Bildung und manuellen Tätigkeiten leiden überproportional. Die Einkommensungleichheit im Alter nimmt zu.

In der zweiten Studie werden die Effekte einer Erhöhung der Altersgrenze für den frühestmöglichen Rentenzugang auf die Sparquoten geschätzt. Theoretisch kann sich ein Rückgang der Großzügigkeit des Rentensystems sowohl positiv als auch negativ auf die Sparquote auswirken. Das Vorzeichen des Effekts hängt von den korrespondierenden Beschäftigungseffekten der Reform ab. Im Jahr 1999 wurde durch eine Reform das frühestmögliche Ren-

tenzugangsalter deutscher Frauen, die nach 1951 geboren wurden, um 3 Jahre angehoben, wodurch eine Diskontinuität entlang der Geburtskohorten entstand. Diese Diskontinuität wird im Rahmen eines Regression Discontinuity Designs verwendet, um den isolierten Effekt des Rentenzugangsalters auf die Sparquoten zu schätzen. Die Schätzung basiert auf detaillierten Einkommens- und Verbrauchsdaten aus einer Haushaltserhebung. Im Gegensatz zur bisherigen Literatur zeigen die Ergebnisse, dass ein Anstieg des frühestmöglichen Rentenzugangsalters keinen positiven Einfluss auf die Sparquote hat. Ein nichtpositiver Effekt steht im Einklang mit der erheblichen Ausweitung der Beschäftigung der betroffenen Kohorten. Insbesondere Haushalte mit geringen Vermögen oder höherer Bildung senken ihre Sparquote gar, wenn sie von der Erhöhung des frühestmöglichen Zugangsalters betroffen sind. Als potenzielle Mechanismen werden ein heterogener Grenznutzen des Konsums und ein Mangel an finanzieller Bildung gering gebildeter Haushalte diskutiert.

In der dritten Studie werden die Auswirkungen des Ruhestands auf die Gesundheit analysiert. Auch hier wird der diskontinuierliche Anstieg des frühestmöglichen Rentenzugangsalters deutscher Frauen in einem Regression Discontinuity Design genutzt. Die Analyse basiert auf einem Two-Stage Least Squares-Schätzverfahren mit zwei unterschiedlichen Mikrodatensätzen. Zum einen werden Surveydaten, zum anderen administrative Daten verwandt. Die Auswirkungen des Ruhestands auf den selbsteingeschätzten Gesundheitszustand werden unter besonderer Berücksichtigung der Effektheterogenität entlang des Bildungsniveaus untersucht. Bei vorsichtiger Interpretation der Schätzergebnisse ist festzuhalten, dass sich der Ruhestand nicht nachteilig auf die Gesundheit auswirkt. Frauen mit niedrigem Bildungsstand zeigen einen positiven Einfluss des Ruhestands auf die Gesundheit.

Die drei Studien zeigen, dass die Auswirkungen von Rentenreformen heterogen sind. In Bezug auf Beschäftigung und Rentenzugangsalter reagieren armutsgefährdete Gruppen weniger stark auf eine Anhebung des Regelrentenalters. In der Folge nimmt die Einkommensungleichheit im Alter zu. Die Analyse der Wirkung des frühestmöglichen Rentenalters deutet auf Probleme von Personen mit niedrigem Bildungsniveau hin, ihre Ersparnisse optimal anzupassen. Darüber hinaus sind gering ausgebildete Menschen der Treiber des positiven Einflusses des Ruhestandes auf die Gesundheit. Daher kann abschließend gesagt werden, dass soziale Gruppen unterschiedlich gut mit Rentenreformen umzugehen wissen, sodass sich gesellschaftliche Ungleichgewichte zu verschlimmern drohen.